

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:59:10 ; Search time 189.33 Seconds  
(without alignments)  
2.528 Million cell updates/sec

Title: US-09-343-406-19

Perfect score: 73

Sequence: 1 ILIKCDERGMIPS 14

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 141170

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	100.0	14	16	R76642
2	73	100.0	14	16	W18850
3	53	72.6	20	16	R72280
4	53	72.6	20	21	Y59553
5	47	64.4	20	16	R72281
6	47	64.4	20	21	Y59554
7	41	56.2	20	16	R86668
8	37	50.7	20	18	W18849
9	34	46.6	20	15	R47505
10	32	43.8	14	16	R76643
11	32	43.8	14	18	W18851
12	31	42.5	23	12	R13416

13	30	41.1	8	20	Y08174
14	29	39.7	15	16	R68602
15	29	39.7	15	16	R68604
16	29	39.7	23	18	W15592
17	29	39.7	23	18	W88049
18	29	39.7	23	21	Y81898
19	28	38.4	11	17	R92106
20	28	38.4	12	14	R41572
21	28	38.4	15	20	Y30504
22	27	37.0	9	14	R41613
23	27	37.0	9	14	R41624
24	27	37.0	10	19	W41465
25	27	37.0	10	19	W47272
26	27	37.0	13	2	P10115
27	27	37.0	13	20	Y43229
28	27	37.0	15	20	Y42762
29	27	37.0	21	21	Y98327
30	27	37.0	22	21	Y98435
31	26	35.6	7	19	Y20455
32	26	35.6	12	18	W47807
33	26	35.6	13	20	W95069
34	26	35.6	17	20	W98997
35	26	35.6	20	18	W40790
36	26	35.6	21	20	Y27760
37	26	35.6	23	21	Y52257
38	26	35.6	23	21	Y52257
39	26	35.6	24	20	W87558
40	25	34.2	8	20	Y08198
41	25	34.2	8	20	Y08173
42	25	34.2	8	20	Y08186
43	25	34.2	9	18	W15693
44	25	34.2	9	20	Y10130
45	25	34.2	10	16	R83761

#### ALIGNMENTS

RESULT 1  
R76642  
ID R76642 standard; peptide; 14 AA.  
AC R76642;  
DE 05-MAR-1996 (first entry)  
DE Peptide derived from human glutamic acid decarboxylase 1.  
KW diabetes; T-cell subpopulation; detection: antigen production;  
KW diagnosis; autoimmune disease.  
XX Homo sapiens.  
XX DE4418091-AL.  
XX PD 27-JUL-1995.  
XX PF 24-MAY-1994; 94DE-4418091.  
XX PR 04-FEB-1994; 94DE-4403522.  
XX PR 20-JAN-1994; 94DE-4401629.  
XX PA (ENDL/) ENDL J.  
XX PA (BOEF) BOEHRINGER MANNHEIM GMBH.  
XX PI Albert W, Dormair K, Endl J, Jung G, Meinel E;  
XX PI Stahl P, Schendel D;  
XX DR WPI, 1995-264505/35.  
XX Antigen-specific activated T-lymphocytes and their detection - by  
XX interaction with inventive peptide(s) of peptide-MHC complexes;  
XX useful in diagnosis of e.g. diabetes and autoimmune diseases

```

XX Claim 1; Fig 1; 21pp; German.
PS
CC R76642-62 are derived from human glutamic acid decarboxylase and
CC specifically react with T-cell sub-populations isolated from recently
CC diagnosed Type-I diabetics. Pharmaceutical compns. contg. these
CC peptides and those shown in R88668-72, are useful for the diagnosis of
CC a disease or predispositions of immune system diseases, tumours, and
CC autoimmune diseases, including diabetes. The peptides are able to detect
CC specific T-cell subpopulations that are then used for antigen prodn.,
CC e.g. by reinjection.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 16; Length 14;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILIKDERGKMIPS 14
   |||||
Db 1 ilikdergkmips 14

RESULT 2
W18850
ID W18850 standard; peptide; 14 AA.
XX
AC W18850;
XX
DT 05-JAN-1998 (first entry)
XX
DE 65 kD glutamic acid decarboxylase peptide fragment 1.
XX
KW GAD: 65 kD; human; glutamic acid decarboxylase; auto-reactive; diagnosis;
KW insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;
KW predisposition; autoimmune; tumour; rheumatoid arthritis;
KW multiple sclerosis.
XX
OS Synthetic.
XX
PN DE19526561-A1.
XX
PD 23-JAN-1997.
XX
PF 20-JUL-1995; 95DE-1026561.
XX
PR 20-JUL-1995; 95DE-1026561.
XX
PA (BOE) BOEHRINGER MANNHEIM GMBH.
XX
PI Donie F, Endl J, Ganz M, Jung G, Kientzsch-engel R;
PI Pozzilli P, Stahl P;
XX
PI WPI; 1997-088254/09.
XX
DR
XX
PT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -
PT involving intradermal admin. of auto-reactive substances
XX
PS Claim 1; Fig 1; 12pp; German.
XX
SQ W18842-70 are peptide fragments of the 65 kD human glutamic acid
CC decarboxylase (GAD). The fragments are auto-reactive substances used for
CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is
CC determined by using a claimed method for diagnosis of cell-mediated
CC diseases or a predisposition to cell-mediated diseases, which is effected
CC by administering an auto-reactive substance intradermally and establishing
CC the diagnosis on the basis of the occurrence or lack of a positive
CC reaction at the site of administration. The method is used for diagnosis
CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases
CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.
XX
SQ Sequence 14 AA;

```

```

Query Match 100.0%; Score 73; DB 18; Length 14;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILIKDERGKMIPS 14
   |||||
Db 1 ilikdergkmips 14

RESULT 3
R72280
ID R72280 standard; peptide; 20 AA.
XX
AC R72280;
XX
DT 13-NOV-1995 (first entry)
XX
DE Glutamic acid decarboxylase (GAD65) fragment.
XX
KW Glutamic acid decarboxylase; GAD65; autoimmune disorders;
KW insulin-dependant diabetes mellitus; stiff man disease.
XX
OS Homo sapiens.
XX
PN W09507992-A.
XX
PD 23-MAR-1995.
XX
PF 24-AUG-1994; 94WO-US09478.
XX
PR 17-SEP-1993; 93US-0123859.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;
XX
PI WPI; 1995-131360/17.
XX
DR
XX
PT New polypeptide fragments of glutamic acid decarboxylase - for
PT diagnosis and treatment of autoimmune disease, esp. insulin
PT dependent diabetes, also related nucleic acid, vectors,
PT antibodies, hydridoma(s) etc.
XX
PS Claim 1; Page 76; 100pp; English.
XX
PX Q86481 and Q86482 encode R71733 and R79105, rat and human glutamic
CC acid decarboxylase (GAD65) respectively, from which the GAD65
CC fragments described in R72261-R72298 were derived. These fragments
CC can be used to detect autoantibodies against GAD, e.g. to diagnose
CC and treat GAD-related autoimmune disorders, such as insulin
CC dependant diabetes mellitus or stiff man disease.
XX
SQ Sequence 20 AA;

Query Match 72.6%; Score 53; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILIKDERGK 10
   |||||
Db 11 ilikdergk 20

RESULT 4
Y59553
ID Y59553 standard; peptide; 20 AA.
XX
AC Y59553;
XX
DT 03-APR-2000 (first entry)

```

XX GAD65 fragment, peptide #20.  
 DE  
 XX  
 KW GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;  
 KW insulin dependent diabetes mellitus; stiff man disease; diagnosis;  
 KW therapy.  
 XX  
 OS Homo sapiens.  
 PN US5998366-A.  
 PD 07-DEC-1999.  
 XX  
 PF 09-APR-1997; 97US-0827618.  
 XX  
 PR 07-JUN-1995; 95US-0485725.  
 PR 21-SEP-1990; 90US-0586536.  
 PR 18-JUN-1991; 91US-0716909.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Tobin AJ, Kaufman DL, Erlander MG;  
 DR WPI: 2000-095930/08.  
 XX  
 PT Ameliorating glutamic acid decarboxylase associated autoimmune  
 PT disorders such as insulin dependent diabetes mellitus and stiff man  
 PT disease -  
 XX  
 PS Claim 1; Column 42; 61pp; English.  
 XX  
 CC This sequence represents a fragment of the glutamic acid decarboxylase 65  
 CC (GAD65) protein. The invention relates to a method of ameliorating GAD  
 CC associated autoimmune disorders by administering a GAD65 peptide to the  
 CC patient. The method can be used for ameliorating GAD associated  
 CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)  
 CC and stiff man disease. GAD65 can also be useful for screening drugs that  
 CC alter GAD function, for generating monoclonal antibodies and in  
 CC immunassays. GAD65 is an effective diagnostic tool for predicting IDDM  
 CC and the diagnosis is quite easy. It is also possible to obtain much  
 CC larger quantities of polypeptide via recombinant techniques than are  
 CC available from natural sources.  
 CC  
 CC Sequence 20 AA:  
 SQ

Query Match 72.6%; Score 53; DB 21; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.0027;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILIKDERGK 10  
 |||||  
 DB 11 lllkcderyk 20

RESULT 5  
 R72281 R72281 standard; Peptide; 20 AA.  
 AC R72281;  
 XX  
 DT 13-NOV-1995 (first entry)  
 XX  
 DE Glutamic acid decarboxylase (GAD65) fragment.  
 XX  
 KW Glutamic acid decarboxylase; GAD65; autoimmune disorder;  
 KW insulin-dependent diabetes mellitus; stiff man disease.  
 XX  
 OS Homo sapiens.  
 PN W09507992-A.  
 PD 23-MAR-1995.

XX  
 PF 24-AUG-1994; 94WO-US09478.  
 XX  
 PR 17-SEP-1993; 93US-0123859.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;  
 DR WPI: 1995-131360/17.  
 XX  
 PT New polypeptide fragments of glutamic acid decarboxylase - for  
 PT diagnosis and treatment of autoimmune disease, esp. insulin  
 PT dependent diabetes, also related nucleic acid, vectors,  
 PT antibodies, hybridoma(s) etc.  
 XX  
 PS Claim 1; Page 76; 100pp; English.  
 XX  
 CC Q86481 and Q86482 encode R71733 and R79105, rat and human glutamic  
 CC acid decarboxylase (GAD65) respectively, from which the GAD65  
 CC fragments described in R72261-R72298 were derived. These fragments  
 CC can be used to detect autoantibodies against GAD, e.g. to diagnose  
 CC and treat GAD-related autoimmune disorders, such as insulin  
 CC dependant diabetes mellitus or stiff man disease.  
 CC  
 CC Sequence 20 AA:  
 SQ

Query Match 64.4%; Score 47; DB 16; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.034;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 DERGNMIPS 14  
 |||||  
 DB 1 dergkmips 9

RESULT 6  
 Y59554 Y59554 standard; peptide; 20 AA.  
 ID Y59554;  
 AC Y59554;  
 XX  
 DT 03-APR-2000 (first entry)  
 XX  
 DE GAD65 fragment, peptide #21.  
 XX  
 KW GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;  
 KW insulin dependent diabetes mellitus; stiff man disease; diagnosis;  
 KW therapy.  
 XX  
 OS Homo sapiens.  
 PN US5998366-A.  
 PD 07-DEC-1999.  
 XX  
 PF 09-APR-1997; 97US-0827618.  
 XX  
 PR 07-JUN-1995; 95US-0485725.  
 PR 21-SEP-1990; 90US-0586536.  
 PR 18-JUN-1991; 91US-0716909.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Tobin AJ, Kaufman DL, Erlander MG;  
 DR WPI: 2000-095930/08.  
 XX  
 PT Ameliorating glutamic acid decarboxylase associated autoimmune  
 PT disorders such as insulin dependent diabetes mellitus and stiff man  
 PT disease -  
 XX

PS Claim 1; Column 42; 61pp; English.

CC This sequence represents a fragment of the glutamic acid decarboxylase 65  
 CC (GAD65) protein. The invention relates to a method of ameliorating GAD  
 CC associated autoimmune disorders by administering a GAD65 peptide to the  
 CC patient. The method can be used for ameliorating GAD associated  
 CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)  
 CC and stiff man disease. GAD65 can also be useful for screening drugs that  
 CC alter GAD function, for generating monoclonal antibodies and in  
 CC immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM  
 CC and the diagnosis is quite easy. It is also possible to obtain much  
 CC larger quantities of polypeptide via recombinant techniques than are  
 CC available from natural sources.

CC Sequence 20 AA;

Query Match 64.4%; Score 47; DB 21; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.034;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 DERGMIPS 14  
 |||||||  
 Db 1 dergmips 9

RESULT 7

R88668 R88668 standard; peptide; 20 AA.

AC R88668;

DT 05-MAR-1996 (first entry)

DE Human glutamic acid decarboxylase residues 306-325.

KW diabetes; T-cell subpopulation; detection; antigen production;  
 diagnosis; autoimmune disease.

XX Homo sapiens.

OS DE4418091-A1.

PN 27-JUL-1995.

PD 24-MAR-1994; 94DE-4418091.

PF 04-FEB-1994; 94DE-4403522.

PR 20-JAN-1994; 94DE-4401629.

XX (ENDL/) ENDL J.  
 (BOEF) BOEHRINGER MANNHEIM GMBH.

PA Albert W, Dormair K, Endl J, Jung G, Meinel E;  
 Stahl P, Schendel D;

PI WPI; 1995-264505/35.

DR WPI; 1995-264505/35.

XX Antigen-specific activated T-lymphocytes and their detection - by  
 PT interaction with inventive peptide(s) of peptide-MHC complexes;  
 PT useful in diagnosis of e.g. diabetes and autoimmune diseases

PS Claim 1; Page 14; 21pp; German.

XX R88668 is derived from residues 306-325 of human glutamic acid  
 CC decarboxylase and specifically reacts with a T-cell sub-population  
 CC isolated from recently diagnosed Type-I diabetics. Pharmaceutical  
 CC compns. consg. this peptide and those shown in R88667 and R76642-62,  
 CC are useful for the diagnosis of a disease or predispositions of immune  
 CC system diseases, tumours, and autoimmune diseases, including diabetes.  
 CC The peptides are able to detect specific T-cell subpopulations that  
 CC are then used for antigen prodn., e.g. by reinjection.

SO Sequence 20 AA;

Query Match 56.2%; Score 41; DB 16; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 ERGMIPS 14  
 |||||||  
 Db 1 ergmips 8

RESULT 8

W18849 W18849 standard; peptide; 20 AA.

AC W18849;

DT 05-JAN-1998 (first entry)

DE 65 kD Glutamic acid decarboxylase peptide fragment VIII.

KW GAD; 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;  
 insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;  
 prediagnosis; autoimmune; tumour; rheumatoid arthritis;  
 multiple sclerosis.

XX Synthetic.

OS DE19526561-A1.

PN 23-JAN-1997.

PD 20-JUL-1995; 95DE-1026561.

PF 20-JUL-1995; 95DE-1026561.

PR 20-JUL-1995; 95DE-1026561.

XX (BOEF) BOEHRINGER MANNHEIM GMBH.

PA Donle F, Endl J, Ganz M, Jung G, Kientsch-engel R;  
 Pozzilli P, Stahl P;

PI WPI; 1997-088254/09.

DR WPI; 1997-088254/09.

XX Skin test for diagnosis of cell-mediated diseases, esp. diabetes -  
 PT involving intradermal admin. of auto-reactive substances

PS Claim 11; Page 10; 12pp; German.

CC W18842-70 are peptide fragments of the 65 kD human glutamic acid  
 CC decarboxylase (GAD). The fragments are autoreactive substances used for  
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is  
 CC determined by using a claimed method for diagnosis of cell-mediated  
 CC diseases or a predisposition to cell-mediated diseases, which is effected  
 CC by administering an autoreactive substance intradermally and establishing  
 CC the diagnosis on the basis of the occurrence or lack of a positive  
 CC reaction at the site of administration. The method is used for diagnosis  
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases  
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.

SO Sequence 20 AA;

Query Match 50.7%; Score 37; DB 18; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 ERGMIP 13  
 |||||||  
 Db 1 ergmip 7

RESULT 9

```
R47505 ID R47505 standard; peptide: 21 AA.  
XX AC  
XX R47505;  
XX  
DT 11-JUL-1994 (first entry)  
XX  
DE Tumour suppressor gene DCC epitope DCCL.  
XX  
KW Epitope; tumour suppressor gene; DCC; monoclonal antibody; Mab; mouse;  
KM KM890; IgG3 subclass; cancer.  
XX  
OS Homo sapiens.  
XX EP577028-A.  
PN  
PD 05-JAN-1994.  
XX  
PE 25-JUN-1993; 93EP-0110213.  
XX  
PR 25-JUN-1992; 92JP-0167101.  
XX  
PA (KYOW ) KYOMA HAKKO KOGYO CO LTD.  
PL Baba S, Hanai N, Ozawa T;  
PI WPI: 1994-009337/02.  
DR  
XX Anti-DCC gene prod. specific monoclonal antibody - useful for  
PT immuno-histochemical staining and cancer diagnosis  
PS Example 1; Page 10; 14pp; English.  
XX  
The sequences given in R47505-09 represent epitopic fragments of  
CC the protein encoded by the tumour suppressor gene DCC. These  
CC peptides are reactive with the monoclonal antibody (Mab) of the  
CC invention, prel. the mouse Mab KM890, belonging to the IgG3 subclass.  
CC This Mab may be used in immunohistochemical staining for  
CC quantitative determination of DCC gene products, as well as for  
CC cancer diagnosis, eg. by histological tissue examination and imaging.  
XX  
Sequence 21 AA:  
SQ
```

```
Query Match             46.6%; Score 34; DB 15; Length 21;  
Best Local Similarity   35.7%; Pred.No. 8.4;  
Matches    5; Conservative      6; Mismatches     3; Indels    0; Gaps    0;  
  
QY       I ILIKCDERGMIPS 14  
          :|::||: |: |: |  
DB        7 VILKCEVIGEMPMT 20  
  
RESULT 10  
R76643 ID R76643 standard; peptide: 14 AA.  
XX AC R76643;  
XX DT  
XX 05-MAR-1996 (first entry)  
XX Peptide derived from human glutamic acid decarboxylase 2.  
XX KW diabetes; T-cell subpopulation; detection; antigen production;  
KM diagnosis; autoimmune disease.  
XX OS Homo sapiens.  
XX DE4418091-AA.  
PN XX  
PD 27-JUL-1995.  
XX  
PF 24-MAY-1994; 94DE-4418091.
```

XX	04-FEB-1994;	94DE-4403522.
PR	20-JAN-1994;	94DE-4401629.
PA	(ENDL/)	ENDL J.
PA	(BOEF )	BOEHRINGER MANNHEIM GMBH.
XX		
PI	Albert W, Dormmair K, Endl J, Jung G, Meinel E;	
PI	Stahl P, Schendel D;	
XX		
DR	WPI; 1995-264505/35.	
XX		
XX	Antigen-specific activated T-lymphocytes and their detection - by	
PT	interaction with inventive peptide(s) of peptide-MHC complexes;	
PT	useful in diagnosis of e.g. diabetes and auto-immune diseases	
XX		
PS	Claim 1; Fig 1; 21pp; German.	
XX		
CC	R76642-62 are derived from human glutamic acid decarboxylase and	
CC	specifically react with T-cell sub-populations isolated from recently	
CC	diagnosed Type-I diabetics. Pharmaceutical compns. confg. these	
CC	peptides and those shown in R7571-72, are useful for the diagnosis of	
CC	a disease or predispositions of immune system diseases, tumours, and	
CC	autoimmune diseases, including diabetes. The peptides are able to detect	
CC	specific T-cell subpopulations that are then used for antigen prodn.,	
CC	e.g. by reinjection.	
XX		
SO	Sequence 14 AA;	
XX		
XX	Query Match	43.8%; Score 32; DB 16; Length 14;
XX	Best Local Similarity	100.0%; Pred. No. 13;
XX	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 ILIKCD 6	
DB	9 ILIKCD 14	
XX		
RESULT 11		
ID	W18851	
ID	W18851 standard; peptide: 14 AA.	
AC	W18851;	
XX		
DT	05-JAN-1998 (first entry)	
XX		
DE	65 KD Glutamic acid decarboxylase peptide fragment 12.	
XX		
XX	GAD; 65 KD; human; glutamic acid decarboxylase; autoreactive; diagnosis;	
XX	Insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;	
KW	predisposition; autoimmune; tumour; rheumatoid arthritis;	
KW	multiple sclerosis.	
XX		
OS	Synthetic.	
XX		
PN	DE19526561-A1.	
XX		
PD	23-JAN-1997.	
XX		
PF	20-JUL-1995; 95DE-1026561.	
XX		
PR	20-JUL-1995; 95DE-1026561.	
XX		
PA	(BOEF ) BOEHRINGER MANNHEIM GMBH.	
XX		
PI	Donle F, Endl J, Ganz M, Jung G, Kientsch-engel R;	
PI	Pozzilli P, Stahl P;	
XX		
DR	WPI; 1997-088254/09.	
XX		
PT	Skin test for diagnosis of cell-mediated diseases, esp. diabetes -	
PT	involving intradermal admn. of auto-reactive substances	



XX MO9428133-A.  
 PN 08-DEC-1994.  
 PD 23-MAY-1994; 94MO-US05769.  
 PF 21-MAY-1993; 93US-0066384.  
 PR (AMGE-) AMGEN INC.  
 PA  
 XX Hu S, Koski RA, Liu N, Pierce GF, Sugarman BJ;  
 PI Wen D;  
 PT WPI; 1995-022805/03.  
 DR New recombinant neu differentiation factors and corresp. DNA -  
 XX are used in the treatment of tumours, dermal wounds, and  
 PT gastrointestinal, kidney and inflammatory bowel diseases.  
 PS Example 14; Page 94; 341pp; English.  
 CC This sequence represents the N-terminal fragment of the human neu  
 CC differentiation factor (NDF), NDF-alpha2(177-241). Recombinantly  
 CC produced NDF bearing this N-terminal sequence was found to stimulate  
 CC tyrosine phosphorylation of the neu receptor protein in an activity  
 CC assay. NDF peptides have the ability to induce a differentiated  
 CC phenotype in certain cell lines and can stimulate or inhibit  
 CC proliferation of certain cell lines. The NDF's can be used to treat  
 CC wounds, tumours derived from epithelial tissue of the breast, stomach  
 CC etc., gastrointestinal disease, Barrett's oesophagus, (non-)cystic  
 CC kidney disease or inflammatory bowel disease. The cDNA sequences  
 CC encoding these protein sequences were used in the isolation of related  
 CC sequences from human cDNA libraries which encode human NDF's.  
 CC  
 SQ Sequence 15 AA;

Query Match 39.7%; Score 29; DB 16; Length 15;  
 Best Local Similarity 55.6%; Pred. No. 48;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 LKCDERCK 10  
 I:| | | | |  
 Db 3 lvkcaek 11

## RESULT 15

R68604 ID R68604 standard; Peptide; 15 AA.

AC R68604;

DT 01-SEP-1995 (first entry)

DE met-NDF-beta(177-246) N-terminal fragment.

XX Alpha; beta; neu differentiation factor; NDF; human; rat; p185-neu;  
 KW tyrosine phosphorylation; differentiation; phenotype; proliferation;  
 KW wound; tumour; epithelial tissue; breast; stomach; PCR; amplify;  
 KW gastrointestinal disease; Barrett's oesophagus; primer;  
 KW (non-)cystic kidney disease; inflammatory bowel disease.

OS Rattus rattus.

PN MO9428133-A.

PD 08-DEC-1994.

PF 23-MAY-1994; 94MO-US05769.

PR 21-MAY-1993; 93US-0066384.

XX

PA (AMGE-) AMGEN INC.  
 XX Hu S, Koski RA, Liu N, Pierce GF, Sugarman BJ;  
 PI Wen D;  
 PT WPI; 1995-022805/03.  
 DR New recombinant neu differentiation factors and corresp. DNA -  
 XX are used in the treatment of tumours, dermal wounds, and  
 PT gastrointestinal, kidney and inflammatory bowel diseases.  
 PS Example 14; Page 100; 341pp; English.  
 CC This sequence represents the N-terminal fragment of the human neu  
 CC differentiation factor (NDF), NDF-beta(177-246). Recombinantly  
 CC produced NDF bearing this N-terminal sequence was found to stimulate  
 CC tyrosine phosphorylation of the neu receptor protein in an activity  
 CC assay. NDF peptides have the ability to induce a differentiated  
 CC phenotype in certain cell lines and can stimulate or inhibit  
 CC proliferation of certain cell lines. The NDF's can be used to treat  
 CC wounds, tumours derived from epithelial tissue of the breast, stomach  
 CC etc., gastrointestinal disease, Barrett's oesophagus, (non-)cystic  
 CC kidney disease or inflammatory bowel disease. The cDNA sequences  
 CC encoding these protein sequences were used in the isolation of related  
 CC sequences from human cDNA libraries which encode human NDF's.  
 CC  
 SQ Sequence 15 AA;

Query Match 39.7%; Score 29; DB 16; Length 15;  
 Best Local Similarity 55.6%; Pred. No. 48;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 LKCDERCK 10  
 I:| | | | |  
 Db 4 lvkcaek 12

## RESULT 16

W11592 ID W11592 standard; peptide; 23 AA.

AC W11592;

DT 03-APR-1997 (first entry)

DE Human breast cancer glycoprotein gp30 tryptic peptide 5.

KW Heparin-binding growth factor; erbB-2 ligand; EGFR ligand;  
 KW epidermal growth factor receptor; tyrosine phosphorylation;  
 KW breast cancer; marker; heregulin.

OS Homo sapiens.

PN US5578482-A.

PD 26-NOV-1996.

PF 25-MAY-1990; 90US-0528438.

PR 26-JUL-1993; 93US-0096277.

PR 25-MAY-1990; 90US-0528438.

PR 14-JAN-1991; 91US-0640497.

PR 22-APR-1992; 92US-0872114.

PR 29-APR-1992; 92US-0875788.

PR 24-JUL-1992; 92US-0917988.

PA (GEOT) UNIV GEORGETOWN.

PI Lipman ME, Lupu R;  
 DR WPI; 1997-020415/02.  
 XX

PT Protein secreted by SK-BR-3 cancer cells - for inhibiting or  
 stimulating growth of cells overexpressing erbB-2  
 XX  
 XX Example 20: Fig 17B; 61pp; English.  
 XX  
 CC Polypeptides were purified from media conditioned by the oestrogen  
 CC receptor negative human breast cancer cell line MDA-MB-231 and the  
 CC SK-BR-3 breast cancer cell line. The polypeptides were partially  
 CC digested with trypsin and the tryptic peptides were separated by  
 CC HPLC on a C18-reversed phase column. Single peaks were subjected to  
 CC protein sequence analysis; five independent fragments were  
 CC sequenced. The present sequence was obtained from peak 5.  
 CC  
 XX  
 XX Sequence 23 AA:  
 SQ  
  
 Query Match 39.7%; Score 29; DB 18; Length 23;  
 Best Local Similarity 55.6%; Pred. No. 76;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 LKCDERGR 10  
 Db 1 LKCAEK 9  
  
 RESULT 17  
 W88049  
 ID W88049 standard; peptide; 23 AA.  
 XX  
 AC W88049;  
 XX  
 DT 12-APR-1999 (first entry)  
 XX  
 DE Amino acid sequence of a gp30 tryptic peptide.  
 XX  
 DE  
 KW GP30: growth factor; erbB-2; erbB-2 receptor;  
 KW epidermal growth factor receptor; EGFR; phosphorylation; antibody;  
 KW erb-2 ligand; marker; tumour progression.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5869618-A.  
 XX  
 PD 09-FEB-1999.  
 XX  
 PF 31-OCT-1995; 95US-0550815.  
 XX  
 PR 26-JUL-1993; 93US-0096277.  
 PR 25-MAY-1990; 90US-0528438.  
 PR 14-JAN-1991; 91US-0640497.  
 PR 22-APR-1992; 92US-0872114.  
 PR 29-APR-1992; 92US-0875788.  
 PR 24-JUL-1992; 92US-0917988.  
 PR 31-OCT-1995; 95US-0550815.  
 XX  
 PA (LIPP/) LIPPMAN M E.  
 PA (LUPU/) LUPU R.  
 XX  
 PI Lippman ME, Lupu R;  
 XX  
 DR WPI, 1999-152878/13.  
 XX  
 PT Antibody specific for 30 kd erbB-2 ligand - useful for detecting  
 PT ligand in cancer prognosis or diagnosis  
 XX  
 PS Example 20: Fig 17B; 61pp; English.  
 XX  
 CC W88045-49 represent the peptide fragments obtained after trypsin  
 CC digest of the protein gp30. Gp30 is a 30 kDa growth factor that  
 CC induces phosphorylation of erbB-2 protein p185 in cells that overexpress  
 CC erbB-2. Induces internalisation of the erbB-2 receptor, stimulates  
 CC growth of erbB-2-overexpressing cells at low concentrations, inhibits  
 CC growth of erbB-2-overexpressing cells at high concentrations, reverses

CC Mab 4D5-dependent inhibition of erbB-2-overexpressing cells, induces  
 CC differentiation of erbB-2-overexpressing cells and stimulates  
 CC phosphorylation of epidermal growth factor receptor (EGFR) in  
 CC EGFR-expressing cells. The specification describes an antibody that  
 CC reacts specifically with the gp30 protein, but does not cross-react with  
 CC TGF-alpha, EGF, amphiregulin or HB-EGF. The antibody can be used to  
 CC detect the 30 kd erbB-2 ligand (gp30) in serum or urine samples as a  
 CC prognostic/diagnostic marker for tumour progression.  
 CC  
 XX  
 XX Sequence 23 AA:  
 SQ  
  
 Query Match 39.7%; Score 29; DB 20; Length 23;  
 Best Local Similarity 55.6%; Pred. No. 76;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 LKCDERGR 10  
 Db 1 LKCAEK 9  
  
 RESULT 18  
 Y81898  
 ID Y81898 standard; peptide; 23 AA.  
 XX  
 AC Y81898;  
 XX  
 DT 22-JUN-2000 (first entry)  
 XX  
 DE Human gp30 protein fragment.  
 XX  
 KW Adenocarcinoma; gp30; gp30/alpha1; heparin-Sepharose binding protein;  
 KW oncogene erbB-2; EGFR; epidermal growth factor receptor; cancer;  
 KW diagnostic marker; tumour progression; breast cancer; prostate cancer;  
 KW therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6040290-A.  
 XX  
 PD 21-MAR-2000.  
 XX  
 PF 26-AUG-1996; 96US-0703089.  
 XX  
 PR 26-JUL-1993; 93US-0096277.  
 PR 25-MAY-1990; 90US-0528438.  
 PR 14-JAN-1991; 91US-0640497.  
 PR 22-APR-1992; 92US-0872114.  
 PR 29-APR-1992; 92US-0875788.  
 PR 24-JUL-1992; 92US-0917988.  
 XX  
 PA (GEOU) UNIV GEORGETOWN.  
 XX  
 PI Lupu R, Lippman ME;  
 XX  
 DR WPI, 2000-270163/23.  
 XX  
 PT Novel growth factor ligand useful for stimulating the growth of normal  
 PT cell and inhibiting the growth of the cells which overexpress the  
 PT oncogene erbB-2 product  
 XX  
 PS Example 20: Fig 17B; 61pp; English.  
 XX  
 CC This sequence represents a fragment of the gp30 protein  
 CC clone gp30/alpha1. The invention relates to a protein (1) that binds to  
 CC heparin-Sepharose, has a molecular weight of 30 kDa (22 kDa after  
 CC N-glycanase digestion), an amino acid sequence of a 22 kDa peptide  
 CC produced by in vitro translation of poly-A RNA from MDA-MB-231 cells,  
 CC and produces the same pattern as the translation product of the  
 CC MDA-MB-231 cells upon hydrolysis with *Staphylococcus aureus* V8 protease  
 CC or elastase. (1) is useful for stimulating the growth of normal cells and  
 CC inhibiting the growth of cells which overexpress the oncogene erbB-2  
 CC product or EGFR (epidermal growth factor receptor), where the cells are



CC preferably adenocarcinoma cells. Monoclonal antibodies to the growth  
 CC factor ligand gp30 are useful as a prognostic or diagnostic marker for  
 CC tumour progression. The protein can be used to treat cancers associated  
 CC with the erbB-2 oncogene product overexpression, including breast,  
 CC ovarian, gastric, lung, prostate, salivary gland and thyroid carcinomas.  
 XX  
 SQ Sequence 23 AA;

Query Match 39.7%; Score 29; DB 21; Length 23;  
 Best Local Similarity 55.6%; Pred. No. 76;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 LINCDEGRK 10  
 | : | | : |  
 Db 1 lvkaekex 9

RESULT 19  
 R92106  
 ID R92106 standard; Peptide; 11 AA.  
 XX  
 AC R92106;

DT 19-MAY-1996 (first entry)  
 XX  
 DE Human RIZ cr2 fragment.

XX Retinoblastoma protein-interacting zinc finger; RIZ;  
 KM cell proliferation; tumour; cancer; neuroblastoma; melanoma;  
 KW diagnosis; therapy; transcription.

OS Homo sapiens.

XX MO9606168-A2.

XX 29-FEB-1996.

PF 18-AUG-1995; 95WO-US10574.

PR 06-MAR-1995; 95US-0399411.

PR 18-AUG-1994; 94US-0292683.

XX (LJOL-) LA JOLLA CANCER RES FOUND.

PA Huang S;

DR WPI; 1996-151371/15.

XX Nucleic acid encoding mammalian retinoblastoma protein-interacting  
 PT zinc finger - used to identify cell proliferation modulating agents  
 PT for treatment of tumours, esp. neuroblastoma and melanoma, also used  
 PT for cancer diagnosis.

PS Claim 19; Page 102; 142pp; English.

XX Active fragments of mammalian retinoblastoma protein-interacting  
 CC zinc finger (RIZ), such as rat RIZ cr2 fragment (R92105) and  
 CC human RIZ cr2 fragment (R92106), bind to retinoblastoma,  
 CC interact with transcription factors and can be involved in  
 CC regulating transcription. They are produced by recombinant DNA  
 CC methods, by peptide synthesis or by enzymatic cleavage of RIZ.  
 CC RIZ or its active fragments (see also R92107-09) can be obt.  
 CC by recombinant DNA methods, and used to identify agents that  
 CC modulate cell proliferation and thereby treat tumour growth or  
 CC neurodegenerative disorders.

SQ Sequence 11 AA;

Query Match 38.4%; Score 28; DB 17; Length 11;  
 Best Local Similarity 66.7%; Pred. No. 53;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 IKCDER 8  
 | : | | : |  
 Db 2 lrckek 7

RESULT 20  
 R41752  
 ID R41752 standard; peptide; 12 AA.  
 XX  
 AC R41752;

DT 18-MAR-1994 (first entry)

DE Conserved NT-4 region 25-36 (xenopus).

XX Nerve growth factor; NGF; brain-derived neurotrophic factor; BDNF;  
 KW neurotrophin; NT; p75NGFR; receptor; transmembrane glycoprotein;  
 KW rat; mouse; human; bovine; guinea pig; chicken;  
 KW xenopus; snake; mutant; modification.

PN MO9318066-A.

PD 16-SEP-1993.

PF 08-MAR-1993; 93WO-SE00201.

PR 06-MAR-1992; 92US-0847369.

PA (MOLI/) MOLINER C F I.  
 (PERS/) PERSSON H B.

PI Persson HB, Moliner CFI;

DR WPI; 1993-303406/38.

XX New mutant forms of neurotrophic factor of increased stability -  
 PT have specific positively charged aminoacid(s) replaced, showing  
 PT more selective binding but equiv. biological activity

XX Disclosure; Fig 1; 70pp; English.

XX Comparison of NGF genes from different species has revealed clusters  
 CC of amino acid residues which are highly conserved across different  
 CC ggs. of vertebrates (rat, mouse, human, bovine, guinea pig and chicken  
 CC NGF: R41747; xenopus NGF: R41748; snake NGF: R41749; rat BDNF:  
 CC R41750; rat NT-3: R41751; and xenopus NT-4: R41752).  
 CC New mutant neurotrophic factor (NF) comprises a wild-type NF having,  
 CC as modification, replacement of one or more positively charged amino  
 CC acids, in the amino acid regions 30-34 or 93-98, with uncharged or  
 CC negatively charged amino acids, so that the mutants have reduced  
 CC ability to bind to p75NGFR (a transmembrane glycoprotein which binds  
 CC NGF with low affinity), c.f. wild-type NF.  
 CC The wild-type NF is pref. NGF, BDNF, NT-3 or NT-4. In the 30-34  
 CC region, the modifications are pref. at Lys32 (or Arg32 in NT)  
 CC and/or Lys34. In the 93-98 region, Lys95 is partic. replaced by Ala.  
 CC Esp. for NT-3, Arg32, His34, Asn93 and/or Asn94; for BDNF, Lys95,  
 CC Lys96 and/or Arg97, and for NT-4 Glu94 and Arg96 are replaced.  
 CC The mutants differ from wild-type NF as regards receptor binding  
 CC affinity and specificity, esp. they can bind to trk receptor but not  
 CC to p75NGFR. Modifications in the 25-36 amino acid region improve  
 CC stability.

SQ Sequence 12 AA;

Query Match 38.4%; Score 28; DB 14; Length 12;  
 Best Local Similarity 57.1%; Pred. No. 58;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 6 DERGKMT 12  
 | : | | : |  
 Db 6 ddrgrktv 12

RESULT 21  
 Y30504 standard; peptide; 15 AA.  
 AC Y30504;  
 XX  
 DT 15-NOV-1999 (first entry)  
 DE Mutant sequence of the major epitopic region 2 of the CD2 protein.  
 KW CD2 protein; rapid mutational analysis method; protein epitope mapping;  
 KW binding domain mapping; binding capacity; anti-CD2 antibody;  
 KW anti-CD4 antibody; ligand binding site study.  
 OS Synthetic.  
 OS Homo sapiens.  
 XX US5955264-A.  
 XX  
 XX 21-SEP-1999.  
 XX  
 XX 11-OCT-1994; 94US-0320663.  
 XX  
 XX 27-FEB-1992; 92US-0842465.  
 XX 15-APR-1988; 88US-0181826.  
 PR 11-OCT-1994; 94US-0320663.  
 XX  
 XX (GEHO ) GEN HOSPITAL CORP.  
 XX  
 XX Peterson A, Seed B;  
 PI  
 DR WPI; 1999-550602/46.  
 XX  
 PT Rapid mutational analysis method for mapping protein epitopes  
 PS  
 XX Disclosure; Page -: 27pp; English.  
 XX  
 CC Y30496-513 represent amino acid substitution variants of a major epitopic  
 CC region 2 of the CD2 protein. The peptides were acquired by mutant  
 CC selection. The CD2 protein and epitopes are used to demonstrate the  
 CC method of the invention. The specification describes a rapid mutational  
 CC analysis method for mapping protein epitopes and binding domains, by  
 CC identifying substitution mutations that result in the loss of binding  
 CC capacity. The method may be used for mapping protein epitopes, antigenic  
 CC domains and binding sites. It has been used for mapping binding sites  
 CC for sixteen anti-CD2 and anti-CD4 monoclonal antibodies. The method is  
 CC especially useful for ligand binding site studies for the design of  
 CC new ligands and drugs.  
 CC note: This sequence does not appear in the specification it was created  
 CC using information provided.  
 CC  
 XX Sequence 15 AA:  
 SQ

Query Match 38.4%; Score 28; DB 20; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 74;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 CDEKGM1 12  
 II::II:  
 Db 5 cdtkgknv 12

RESULT 22  
 R41613  
 ID R41613 standard; protein; 9 AA.  
 AC R41613;  
 XX  
 XX 25-MAR-1994 (first entry)  
 DT  
 XX

DE TRIP-32.  
 XX  
 KW Erabutoxin; thrombin; receptor; inhibitor; peptide; TRIP; septicemia;  
 KW fusion protein; plasmid; pIII; phagemid vector; pTZ19R; syngene;  
 KW fibrinogen; coagulation; platelet activation; reperfusion damage;  
 KW neurodegenerative disease; bone resorption; osteoporosis; thrombosis;  
 KW inflammation; adult respiratory distress syndrome; septic shock.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH MISC-difference 4 /note="D-form residue"  
 FT  
 XX  
 XX W09318141-A.  
 XX  
 XX 16-SEP-1993.  
 XX  
 XX 02-MAR-1993; 93WO-US01901.  
 XX  
 XX 02-MAR-1992; 92US-0847561.  
 XX  
 XX (BIOJ ) BIOGEN INC.  
 XX  
 XX Freilinger AL, Maraganore JM;  
 PI  
 DR WPI; 1993-303456/38.  
 XX  
 XX  
 PT New thrombin receptor antagonist peptide(s) and fusion proteins -  
 PT used to inhibit thrombin-induced platelet aggregation and smooth  
 PT muscle cell proliferation  
 XX  
 XX Example 37; Page 46; 95pp; English.  
 XX  
 CC The sequences given in R41583-633 are thrombin receptor inhibitor  
 CC peptides (TRIPs). TRIPs competitively bind to and inhibit the  
 CC thrombin receptor without affecting thrombin activities toward  
 CC fibrinogen and other proteins involved in coagulation. They can  
 CC be used for inhibiting thrombin-induced platelet activation and  
 CC inhibiting thrombin-induced smooth muscle cell proliferation. They  
 CC can be used for treating and preventing neurodegenerative disease,  
 CC treating and preventing bone resorption (eg. in osteoporosis),  
 CC treating and preventing thrombin-induced inflammation, eg. in adult  
 CC respiratory distress syndrome, septic shock, septicemia and  
 CC reperfusion damage) and for inhibiting platelet-dependent thrombosis.  
 CC Fusion proteins containing these sequences may also act as thrombin  
 CC receptor antagonists.  
 CC  
 XX Sequence 9 AA:  
 SQ

Query Match 37.0%; Score 27; DB 14; Length 9;  
 Best Local Similarity 33.3%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 LKCDERK 10  
 I::I:::  
 Db 1 lvcdknsr 9

RESULT 23  
 R41624  
 ID R41624 standard; protein; 9 AA.  
 AC R41624;  
 XX  
 XX 25-MAR-1994 (first entry)  
 DT  
 XX  
 XX TRIP-43.  
 DE  
 XX  
 XX Erabutoxin; thrombin; receptor; inhibitor; peptide; TRIP; septicemia;  
 KW fusion protein; plasmid; pIII; phagemid vector; pTZ19R; syngene;  
 KW fibrinogen; coagulation; platelet activation; reperfusion damage;



XX 22-MAY-1996; 96US-0651650.  
PR (STRD ) UNIV LELAND STANFORD JUNIOR.  
XX  
PA Clayberger C, Krensky AM;  
XX  
PI WPI; 1998-018220/02.  
XX  
DR  
XX  
FT Novel immunomodulatory peptide-type compound - useful for inhibiting  
PT transplant rejection  
XX  
PS Claim 10; Page 36; 41pp: English.  
XX  
CC The present sequence is an immunomodulatory peptide, which  
CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used  
CC in a pharmaceutical composition together with a subtherapeutic dose  
CC of an immunosuppressant, to extend the period of acceptance of a  
CC transplant from a major histocompatibility complex (MHC) unmatched  
CC donor, i.e. to inhibit transplant rejection. It can also be used in  
CC the treatment of autoimmune diseases.  
CC Peptides using the D-form amino acids are more effective  
CC Immunomodulators than their diastereomers or enantiomers.  
XX  
SQ Sequence 10 AA;

Query Match 37.0%; Score 27; DB 19; Length 10;  
Best Local Similarity 62.5%; Pred. No. 73;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ILIKCDE 8  
:11:111  
Db 3 111rider 10

Search completed: February 5, 2001, 10:59:11  
Job time: 634 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:59:10 ; Search time 189.33 seconds

(without alignments)  
2.528 Million cell updates/sec

Title: US-09-343-406-19

Perfect score: 73  
Sequence: 1 ILKCDRCRKMPIS 14

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 141170

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq\_36:\*

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:\*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:\*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:\*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:\*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:\*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:\*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:\*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:\*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:\*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:\*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:\*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:\*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:\*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:\*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:\*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:\*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:\*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:\*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:\*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:\*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:\*

Fred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	73	100.0	14 16 R76642	Peptide derived fr
2	73	100.0	14 18 W18850	65 KD Glutamic aci
3	53	72.6	20 16 R72380	Glutamic acid deca
4	53	72.6	20 21 R59553	GAD65 fragment, pe
5	47	64.4	20 16 R72281	Glutamic acid deca
6	47	64.4	20 21 R59554	GAD65 fragment, pe
7	41	56.2	20 16 R88668	Human glutamic aci
8	37	50.7	20 18 W18849	65 KD Glutamic aci
9	34	46.6	21 15 R47505	Tumour suppressor g
10	32	43.8	14 16 R76643	Peptide derived fr
11	32	43.8	14 18 W18851	65 KD Glutamic aci
12	31	42.5	23 12 R13416	Cardiacstatin-like p

13	30	41.1	8 20 Y08174
14	29	39.7	15 16 R68602
15	29	39.7	15 16 R68604
16	29	39.7	23 18 W15592
17	29	39.7	23 18 W88049
18	29	39.7	23 21 Y81898
19	28	38.4	11 17 R92106
20	28	38.4	12 14 R41752
21	28	38.4	15 20 Y30504
22	27	37.0	9 14 R41613
23	27	37.0	9 14 R41624
24	27	37.0	10 19 W41465
25	27	37.0	10 19 W47272
26	27	37.0	13 2 P10115
27	27	37.0	13 20 Y43229
28	27	37.0	15 20 Y42762
29	27	37.0	21 21 Y98327
30	27	37.0	21 21 Y98327
31	26	35.6	7 19 Y20455
32	26	35.6	12 18 W47807
33	26	35.6	13 20 W95069
34	26	35.6	17 20 W98997
35	26	35.6	20 18 W40790
36	26	35.6	21 20 Y27760
37	26	35.6	23 21 Y52256
38	26	35.6	23 21 Y52257
39	26	35.6	23 21 Y52257
40	25	34.2	8 20 W87558
41	25	34.2	8 20 Y08198
42	25	34.2	8 20 Y08173
43	25	34.2	8 20 Y08186
44	25	34.2	9 18 W15693
45	25	34.2	9 20 Y10130
			R83761

#### ALIGNMENTS

RESULT 1  
R76642  
ID R76642 standard; peptide; 14 AA.  
XX AC R76642;  
XX 05-MAR-1996 (first entry)  
XX  
XX  
XX Peptide derived from human glutamic acid decarboxylase 1.  
DE  
XX  
XX diabetes; T-cell subpopulation; detection; antigen production;  
KW diagnosis; autoimmune disease.  
XX  
XX Homo sapiens.  
XX  
XX DE4418091-A1.  
XX  
XX 27-JUL-1995.  
XX  
XX 24-MAY-1994; 94DE-4418091.  
XX  
XX 04-FEB-1994; 94DE-4403522.  
XX  
XX 20-JAN-1994; 94DE-4401629.  
XX  
XX (ENDL/) ENDL J.  
XX (BOEF) BOEHRINGER MANNHEIM GMBH.  
XX Albert W, Dormair K, Endl J, Jung G, Meinel E;  
XX Stahl P, Schendel D;  
XX WPI; 1995-264505/35.  
XX  
XX Antigen-specific activated T-lymphocytes and their detection - by  
XX interaction with invariant peptide(s) of peptide-MHC complexes;  
XX useful in diagnosis of e.g. diabetes and autoimmune diseases

Clotting factor VI  
NFE-alpha2(177-241  
met-NFE-beta1(177-  
Human breast cance  
Amino acid sequenc  
Human gp30 protein  
Human R12 cr2 frag  
Conserved NF-4 reg  
Mutant sequence of  
TRIP-32. Synthet  
TRIP-43. Synthet  
Fragment of hemato  
Immunomodulatory p  
Sequence which cor  
RGP-containing pep  
Alpha D peptide de  
Alpha D peptide de  
Human microtubule  
Cyclic peptide or  
Acetylcholine-bind  
I domain peptide p  
Peptide which bind  
Human secreted pro  
H. influenzae y9jd  
P. haemolytica y9j  
Consensus zinc fin  
Clotting factor VI  
Clotting factor VI  
Clotting factor VI  
Clotting factor VI  
Platelet aggregati  
T cell epitope/MHC  
NFE derived struct



XX Claim 1; Fig 1; 21pp; German.  
 PS  
 CC R76642-62 are derived from human glutamic acid decarboxylase and  
 CC specifically react with T-cell sub-populations isolated from recently  
 CC diagnosed type-1 diabetics. Pharmaceutical compans. contg. these  
 CC peptides and those shown in R88668-72, are useful for the diagnosis of  
 CC a disease or predispositions of immune system diseases, tumours, and  
 CC autoimmune diseases, including diabetes. The peptides are able to detect  
 CC specific T-cell subpopulations that are then used for antigen prodn.,  
 CC e.g. by reinjection.  
 CC  
 XX Sequence 14 AA;  
 SQ

Query Match 100.0%; Score 73; DB 16; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 4,1e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILIKDERGKMIPS 14  
 Db 1 11lkcderygmips 14  
 |||||  
 |||||

RESULT 2  
 ID W1850 standard; peptide; 14 AA.  
 XX  
 AC W1850;  
 XX  
 DT 05-JAN-1998 (first entry)  
 XX  
 DE 65 kD glutamic acid decarboxylase peptide fragment 1.  
 XX  
 KM GAD; 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;  
 KM insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;  
 KM predisposition; autoimmune; tumour; rheumatoid arthritis;  
 KM multiple sclerosis.  
 XX  
 OS Synthetic.  
 XX  
 PN DEL9526561-AL.  
 XX  
 PD 23-JAN-1997.  
 XX  
 PF 20-JUL-1995; 95DE-1026561.  
 XX  
 PR 20-JUL-1995; 95DE-1026561.  
 XX  
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.  
 XX  
 PI Donle F, Endl J, Ganz M, Jung G, Kientsch-engel R;  
 PI Pozzilli P, Stahl P;  
 XX  
 DR WPI; 1997-088254/09.  
 XX  
 FT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -  
 FT involving intradermal admin. of auto:reactive substances  
 XX  
 PS Claim 11; Fig 1; 12pp; German.  
 XX

W18842-70 are peptide fragments of the 65 kD human glutamic acid  
 CC decarboxylase (GAD). The fragments are autoreactive substances used for  
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is  
 CC determined by using a claimed method for diagnosis of cell-mediated  
 CC diseases or a predisposition to cell-mediated diseases, which is effected  
 CC by administering an autoreactive substance intradermally and establishing  
 CC the diagnosis on the basis of the occurrence or lack of a positive  
 CC reaction at the site of administration. The method is used for diagnosis  
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases  
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.  
 CC  
 XX Sequence 14 AA;  
 SQ

Query Match 100.0%; Score 73; DB 18; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 4,1e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILIKDERGKMIPS 14  
 Db 1 11lkcderygmips 14  
 |||||  
 |||||

RESULT 3  
 ID R72280 standard; Peptide; 20 AA.  
 XX  
 AC R72280;  
 XX  
 DT 13-NOV-1995 (first entry)  
 XX  
 DE Glutamic acid decarboxylase (GAD65) fragment.  
 XX  
 KM Glutamic acid decarboxylase; GAD65; autoimmune disorders;  
 KM insulin-dependant diabetes mellitus; stiff man disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09507992-A.  
 XX  
 PD 23-MAR-1995.  
 XX  
 PF 24-AUG-1994; 94MO-US09478.  
 XX  
 PR 17-SEP-1993; 93US-0123859.  
 XX  
 PA (REGC) UNIV CALIFORNIA.  
 XX  
 PI Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;  
 XX  
 DR WPI; 1995-131360/17.  
 XX  
 PT New polypeptide fragments of glutamic acid decarboxylase - for  
 PT diagnosis and treatment of autoimmune disease, esp. insulin  
 PT dependent diabetes; also related nucleic acid, vectors,  
 PT antibodies, hydridoma(s) etc.  
 XX  
 PS Claim 1; Page 76; 100pp; English.  
 XX  
 CC O86481 and O86482 encode R71733 and R79105, rat and human glutamic  
 CC acid decarboxylase (GAD65) respectively, from which the GAD65  
 CC fragments described in R72261-R72298 were derived. These fragments  
 CC can be used to detect autoantibodies against GAD, e.g. to diagnose  
 CC and treat GAD-related autoimmune disorders, such as insulin  
 CC dependant diabetes mellitus or stiff man disease.  
 XX  
 SQ Sequence 20 AA;  
 SQ

Query Match 72.6%; Score 53; DB 16; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.0027;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILIKDERGK 10  
 Db 11 11lkcderygk 20  
 |||||  
 |||||

RESULT 4  
 ID Y59553 standard; peptide; 20 AA.  
 XX  
 AC Y59553;  
 XX  
 DT 03-APR-2000 (first entry)  
 XX









XX Claim 1; Fig 1; 21pp; German.  
 PS  
 CC R76642-62 are derived from human glutamic acid decarboxylase and  
 CC specifically react with T-cell sub-populations isolated from recently  
 CC diagnosed Type-1 diabetics. Pharmaceutical compans. contg. these  
 CC peptides and those shown in R8668-72, are useful for the diagnosis of  
 CC diseases or predispositions of immune system diseases, tumours, and  
 CC autoimmune diseases, including diabetes. The peptides are able to detect  
 CC specific T-cell subpopulations that are then used for antigen prodn.,  
 CC e.g. by reinjection.  
 CC  
 SQ Sequence 14 AA;  
 Query Match 100.0%; Score 73; DB 16; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ILIKDERGKMPIS 14  
 Db 1 ILIKDERGKMPIS 14  
 RESULT 2  
 W18850  
 ID W18850 standard; peptide; 14 AA.  
 AC W18850;  
 XX  
 XX W18850;  
 DT 05-JAN-1998 (first entry)  
 XX  
 DE 65 kD Glutamic acid decarboxylase peptide fragment 1.  
 XX  
 XX GAD; 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;  
 KW insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;  
 KM predisposition; autoimmune; tumour; rheumatoid arthritis;  
 KM multiple sclerosis.  
 XX  
 OS Synthetic.  
 XX  
 PN DE19526561-A1.  
 XX  
 PD 23-JAN-1997.  
 XX  
 PF 20-JUL-1995; 95DE-1026561.  
 XX  
 PR 20-JUL-1995; 95DE-1026561.  
 XX  
 PA (BOE) BOEHRINGER MANNHEIM GMBH.  
 XX  
 PI Donle F, Endl J, Ganz M, Jung G, Kientsch-engel R;  
 PI Poszilli P, Stahl P;  
 XX  
 DR WPI; 1997-088254/09.  
 XX  
 PT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -  
 PT involving intradermal admin. of auto-reactive substances  
 XX  
 PS Claim 11; Fig 1; 12pp; German.  
 XX  
 CC W18842-70 are peptide fragments of the 65 kD human glutamic acid  
 CC decarboxylase (GAD). The fragments are autoreactive substances used for  
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is  
 CC determined by using a claimed method for diagnosis of cell-mediated  
 CC diseases or a predisposition to cell-mediated diseases, which is effected  
 CC by administering an autoreactive substance intradermally and establishing  
 CC the diagnosis on the basis of the occurrence or lack of a positive  
 CC reaction at the site of administration. The method is used for diagnosis  
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases  
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.  
 CC  
 SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 18; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ILIKDERGKMPIS 14  
 Db 1 ILIKDERGKMPIS 14  
 RESULT 3  
 R72280  
 ID R72280 standard; peptide; 20 AA.  
 AC R72280;  
 XX  
 XX R72280;  
 DT 13-NOV-1995 (first entry)  
 XX  
 DE Glutamic acid decarboxylase (GAD65) fragment.  
 XX  
 XX Glutamic acid decarboxylase; GAD65; autoimmune disorders;  
 KW insulin-dependant diabetes mellitus; stiff man disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09507992-A.  
 XX  
 PD 23-MAR-1995.  
 XX  
 PF 24-AUG-1994; 94WO-US094478.  
 XX  
 PR 17-SEP-1993; 93US-0123859.  
 XX  
 PA (REGC) UNIV CALIFORNIA.  
 XX  
 PI Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;  
 PI WPI; 1995-131360/17.  
 DR  
 XX  
 XX New polypeptide fragments of glutamic acid decarboxylase - for  
 PT diagnosis and treatment of autoimmune disease, esp. insulin  
 PT dependent diabetes, also related nucleic acid, vectors,  
 PT antibodies, hybridoma(s) etc.  
 XX  
 PS Claim 1; Page 76; 100pp; English.  
 XX  
 CC Q86481 and Q86482 encode R71733 and R79105, rat and human glutamic  
 CC acid decarboxylase (GAD65) respectively, from which the GAD65  
 CC fragments described in R72261-R72298 were derived. These fragments  
 CC can be used to detect autoantibodies against GAD, e.g. to diagnose  
 CC and treat GAD-related autoimmune disorders, such as insulin  
 CC dependant diabetes mellitus or stiff man disease.  
 CC  
 SQ Sequence 20 AA;  
 Query Match 72.6%; Score 53; DB 16; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.0027;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ILIKDERGK 10  
 Db 1 ILIKDERGK 20  
 RESULT 4  
 Y59553  
 ID Y59553 standard; peptide; 20 AA.  
 AC Y59553;  
 XX  
 DT 03-APR-2000 (first entry)



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 11:27:24 ; Search time 265.72 Seconds  
(without alignments)  
6.175 Million cell updates/sec

Title: US-09-343-406-19  
Perfect score: 73  
Sequence: 1 ILIKDERGKMIPS 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_15:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	49.3	22	11	088222
2	31	42.5	17	2	052748
3	29	39.7	15	6	09TRN4
4	26	35.6	20	11	090021
5	26	35.6	21	10	041580
6	26	35.6	21	11	088231
7	26	35.6	24	4	09UT67
8	26	35.6	24	10	09S883
9	25	34.2	11	7	029831
10	25	34.2	15	6	09TRP0
11	25	34.2	24	5	P82176
12	24	32.9	20	4	060696
13	24	32.9	25	6	077604
14	24	32.9	25	6	077831
15	24	32.9	25	6	077832
16	23	31.5	11	5	09NR65
17	23	31.5	21	5	09TW67
18	23	31.5	23	4	014347
19	23	31.5	23	13	09PRT7

20	23	31.5	24	2	09R573	Q91573 nitrosomona
21	23	31.5	24	5	094373	Q94373 caenorhabdi
22	23	31.5	25	11	061461	061461 mus musculu
23	23	30.1	15	6	09TRN8	09trn8 sus scrofa
24	22	30.1	15	6	09TRN3	09trn3 sus scrofa
25	22	30.1	16	7	077922	077922 oreochromis
26	22	30.1	19	6	09TV19	09tv19 macaca mula
27	22	30.1	20	2	P97146	P97146 escherichia
28	22	30.1	25	4	09UED3	09ued3 homo sapien
29	22	30.1	25	6	018918	018918 sus scrofa
30	21.5	29.5	16	3	P79034	P79034 emeritella
31	21	28.8	12	4	Q9UMR0	Q9umr0 homo sapien
32	21	28.8	14	2	Q55326	Q55326 synechococc
33	21	28.8	15	4	Q15344	Q15344 homo sapien
34	21	28.8	15	6	09TRP2	09trp2 sus scrofa
35	21	28.8	18	9	Q38571	Q38571 bacteriophag
36	21	28.8	20	6	09TRH8	09trh8 oryctolagus
37	21	28.8	20	10	Q9S8H1	Q9s8h1 hordeum vul
38	21	28.8	20	12	Q41541	Q41541 human immun
39	21	28.8	20	12	Q41542	Q41542 human immun
40	21	28.8	22	1	P71537	P71537 methanosarc
41	21	28.8	22	2	Q85524	Q85524 chlamydia t
42	21	28.8	25	2	Q55093	Q55093 synechocyst
43	21	28.8	25	6	077602	077602 papio anubi
44	21	28.8	25	6	077603	077603 theropithec
45	21	28.8	25	6	077605	077605 mandrillus

#### ALIGNMENTS

RESULT 1

ID 088222 PRELIMINARY; PRT; 22 AA.

AC 088222:

DT 01-NOV-1998 (TREMBLrel\_08, Created)

DT 01-NOV-1998 (TREMBLrel\_08, Last sequence update)

DT 01-JUN-2000 (TREMBLrel\_14, Last annotation update)

DE MS2F72 (FRAGMENT).

GN MS2F72.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_Taxid:10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL6;

RX MEDLINE=98296253; PubMed=9630514;

RA Agata Y., Matsuda E., Shimizu A.;

RT "Rapid and efficient cloning of cDNAs encoding Kruppel-like zinc

RT finger proteins by degenerate PCR."

RL Gene 213:35-64(1998).

DR EMBL; AB010317; BAA31373.1; -.

DR INTERPRO; IPR000822; -.

FT NON\_TER

FT NON\_TER

SQ SEQUENCE 22 AA: 2530 MW: 0CBEC69A2066141A CRC64;

Query Match 49.3%; Score 36; DB 11; Length 22;  
Best Local Similarity 60.0%; Pred. No. 10;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 KCDGKMIIP 13  
111:111

DB 5 KCDGKMIIP 14

RESULT 2

ID 052748 PRELIMINARY; PRT; 17 AA.

AC 052748:

DT 01-JUN-1998 (TREMBLrel\_06, Created)

DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last annotation update)  
 DE HENEROCYST-INHIBITING SIGNALING PEPTIDE.  
 CN PATS.  
 OS Anabaena sp. (strain PCC 7120).  
 CC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.  
 OX NCBI\_TaxID=1174;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PCC 7120;  
 RA Xoon H.-S., Golden J.W.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF046871; AAC03103.1;  
 SQ SEQUENCE 17 AA; 1927 MW; 192BE168476867F3 CRC64;

Query Match 42.5%; Score 31; DB 2; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CDERG 9  
 Db 10 CDERG 14

RESULT 3  
 Q9TRN4 PRELIMINARY; PRT; 15 AA.  
 AC Q9TRN4;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE LOW MR ZONA PELLUCIDA BINDING PROTEIN (FRAGMENT).  
 OS Sus scrofa (pig).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 OX NCBI\_TaxID=96823;  
 RN [1]  
 RP SEQUENCE.  
 RA MEDLINE=92378826; PubMed=1510840;  
 RA Parry R.V., Barker P.J., Jones R.;  
 RT "Characterization of low Mr zona pellucida binding proteins from boar  
 spermatozoa and seminal plasma."  
 RL Mol. Reprod. Dev. 33:108-115(1992).  
 SQ SEQUENCE 15 AA; 1768 MW; 7FA2379E1461610E CRC64;

Query Match 39.7%; Score 29; DB 6; Length 15;  
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 DERGMT 12  
 Db 7 DERGMT 13

RESULT 4  
 Q9OUZ1 PRELIMINARY; PRT; 20 AA.  
 AC Q9OUZ1;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE ALPHA CLASS GLUTATHIONE S-TRANSFERASE SUBUNIT 2 (EC 2.5.1.18)  
 DE (FRAGMENT).  
 OS Rattus sp.  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10118;  
 RN [1]  
 RP SEQUENCE.  
 RA MEDLINE=96036981; PubMed=7485987;  
 RA Rouimi P., Debrauwer L., Tulliez J.;

RT "Electrospray ionization-mass spectrometry as a tool for  
 RT characterization of glutathione S-transferase isozymes."  
 RL Anal. Biochem. 229:304-312(1995).  
 DR HSSP; P24472; IGUK.  
 SQ SEQUENCE 20 AA; 2412 MW; 9498DE2C030B236 CRC64;

Query Match 35.6%; Score 26; DB 11; Length 20;  
 Best Local Similarity 62.5%; Pred. No. 5.6e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 DERGMT 13  
 Db 10 DERGMT 17

RESULT 5  
 Q41580 PRELIMINARY; PRT; 21 AA.  
 AC Q41580;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE RIBULOSEBISPHOSPHATE CARBOXYLASE SMALL SUBUNIT (FRAGMENT).  
 OS Triticum aestivum (wheat).  
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Magnoliophyta; Liliopsida; Poales; Poaceae; Triticum.  
 OX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SEEDLING LEAF;  
 RX MEDLINE=84169511; PubMed=6324097;  
 RA Smith S.W., Bedbrook J.R., Speirs J.;  
 RT "Characterisation of three cDNA clones encoding different mRNAs for  
 RT the precursor to the small subunit of wheat ribulosebiphosphate  
 RT carboxylase."  
 RL Nucleic Acids Res. 11:8719-8734(1983).  
 DR EMBL; K02324; AAA34300.2; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 21 AA; 2226 MW; FAD28BDB9FD41520 CRC64;

Query Match 35.6%; Score 26; DB 10; Length 21;  
 Best Local Similarity 66.7%; Pred. No. 5.9e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 CDERGK 10  
 Db 15 CDERGK 20

RESULT 6  
 O88231 PRELIMINARY; PRT; 21 AA.  
 AC O88231;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE MSZF80-1 (FRAGMENT).  
 GN MSZF80-1.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL6;  
 RX MEDLINE=98296253; PubMed=9630514;  
 RA Agata Y., Matsuda E., Shimizu A.;  
 RT "Rapid and efficient cloning of cDNAs encoding Kruppel-like zinc  
 RT finger proteins by degenerate PCR."  
 RL Gene 213:55-64(1998).  
 DR EMBL; AB010326; BAA31382.1; -.

DR INTERPRO: IPR000822;  
 FT NON\_TER 1  
 FT NON\_TER 21  
 SQ SEQUENCE 21 AA; 2447 MW; 2956C0247613DF43 CRC64;

Query Match 35.6%; Score 26; DB 11; Length 21;  
 Best Local Similarity 71.4%; Pred. No. 5.9e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CDEGRK 10  
 DB 5 KCPECGK 11

RESULT 7  
 ID Q9UE7 PRELIMINARY; PRT; 24 AA.  
 AC Q9UE7;  
 DT 01-MAY-2000 (TRENBLREL. 13, Created)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last annotation update)  
 DE D620E11.1D (NOVEL HELICASE C-TERMINAL DOMAIN AND SNF2 N-TERMINAL  
 GN D620E11.1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sehra H.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL031667; CAB57840.1;  
 FT NON\_TER 1  
 FT NON\_TER 24  
 SQ SEQUENCE 24 AA; 2765 MW; 23DAFF242E5A9D56 CRC64;

Query Match 35.6%; Score 26; DB 4; Length 24;  
 Best Local Similarity 50.0%; Pred. No. 6.7e+02;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 4 CDEGRK 11  
 DB 8 RKEPGK 15

RESULT 8  
 ID Q9S883 PRELIMINARY; PRT; 24 AA.  
 AC Q9S883;  
 DT 01-MAY-2000 (TRENBLREL. 13, Created)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
 DT 01-JUN-2000 (TRENBLREL. 14, Last annotation update)  
 DE ATP SYNTHASE DELTA SUBUNIT (FRAGMENT).  
 OS Chlamydomonas reinhardtii.  
 CC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 OC Chlamydomonadaceae; Chlamydomonas.  
 OX NCBI\_TaxID=3055;  
 RN [1]  
 RP SEQUENCE.  
 RA MEDLINE=96128220; PubMed=8543042;  
 RX Fiedler H.R., Schmid R., Leu S., Shavit N., Strotmann H.,  
 RT "Isolation of CFRCF1 from Chlamydomonas reinhardtii cw15 and the N-  
 RL terminal amino acid sequences of the CFRCF1 subunits.";  
 SQ FEBS Lett. 377:163-166(1995).  
 SEQUENCE 24 AA; 2707 MW; 5CCAE3310F50FB44 CRC64;

Query Match 35.6%; Score 26; DB 10; Length 24;  
 Best Local Similarity 66.7%; Pred. No. 6.7e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 6 DERGKM 11  
 DB 18 DERGKL 23

RESULT 9  
 ID Q29831 PRELIMINARY; PRT; 11 AA.  
 AC Q29831;  
 DT 01-NOV-1996 (TRENBLREL. 01, Created)  
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last annotation update)  
 DE HISTOCOMPATIBILITY ANTIGEN HLA-DR ALPHA (FRAGMENT).  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83169718; PubMed=6403940;  
 RA Das H.K., Bito P.A., Cohen S.N., Erlich H.A., Von Gabala A.,  
 RA Lawrence S.K., Lemaux P.G., McDevitt H.O., Peterlin B.M., Schulz M.F.,  
 RA Sood A.K., Weissman S.M.;  
 RT "Use of synthetic oligonucleotide probes complementary to genes for  
 RT human HLA-DR alpha and beta as extension primers for the isolation of  
 RT 5'-specific genomic clones."  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:1531-1535(1983).  
 DR EMBL; V00525; CAA23784.1;  
 KM MHC.  
 FT NON\_TER 1  
 FT NON\_TER 11  
 SQ SEQUENCE 11 AA; 1230 MW; 9378714E0865B1EA CRC64;

Query Match 34.2%; Score 25; DB 7; Length 11;  
 Best Local Similarity 37.5%; Pred. No. 4.7e+02;  
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 5 CDEGRK 12  
 DB 3 CDHSGRVL 10

RESULT 10  
 ID Q9TRP0 PRELIMINARY; PRT; 15 AA.  
 AC Q9TRP0;  
 DT 01-MAY-2000 (TRENBLREL. 13, Created)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
 DT 01-JUN-2000 (TRENBLREL. 14, Last annotation update)  
 DE LOW MR ZONA PELLUCIDA BINDING PROTEIN (FRAGMENT).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 OX NCBI\_TaxID=98823;  
 RN [1]  
 RP SEQUENCE.  
 RA MEDLINE=92378826; PubMed=1510840;  
 RX Parry R.V., Barker P.J., Jones R.;  
 RT "Characterization of low Mr zona pellucida binding proteins from boar  
 RT spermatozoa and seminal plasma."  
 RL Mol. Reprod. Dev. 33:108-115(1992).  
 SQ SEQUENCE 15 AA; 1525 MW; CC800E1028CF2058 CRC64;

Query Match 34.2%; Score 25; DB 6; Length 15;  
 Best Local Similarity 62.5%; Pred. No. 6.4e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 ERGKMIPS 14  
 DB 4 KRGAIPIS 11

## RESULT 11

PRELIMINARY: PRT: 24 AA.  
 ID P82176  
 AC P82176  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DE METALLOPROTEINASE INHIBITOR (IMPI) (FRAGMENT).  
 OS Galliera mellonella (Wax moth).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Pyraloidea; Pyralidae; Galleriinae; Galleria.  
 NCBI\_TaxID=7137;  
 [1]  
 RN SEQUENCE.  
 RP TISSUE-LARVAL HEMOLYMPH;  
 RC MEDLINE=98409271; PubMed=9738891;  
 RA Wedde M., Weise C., Kopacek P., Franke P., Vilcinskas A.;  
 RT "Purification and characterization of an inducible metalloprotease inhibitor from the hemolymph of greater wax moth larvae, Galliera mellonella.";  
 RL Eur. J. Biochem. 255:535-543(1998).  
 CC -1- FUNCTION: IT INHIBITS THERMOLYSIN AND THE METALLOPROTEASE OF B. FOLIOYMYXA. NO ACTIVITY ON TRIPSPIN OR CISTEINPROTEASE PAPAIN.  
 CC -1- INDUCTION: BY LIPOPOLYSACCHARIDE (LPS).  
 CC -1- PTM: FIVE DISULFIDE BONDS ARE PRESENT.  
 CC -1- PTM: THE HEAVY CHAIN MAY BE O-GLYCOSYLATED.  
 CC -1- MASS SPECTROMETRY: MW=8360; METHOD=MALDI.  
 KW Metalloprotease inhibitor; Glycoprotein.  
 FT NON\_TER 1 24 1  
 FT NON\_TER 24 24 1  
 SQ SEQUENCE 24 AA; 2783 MW; 490E278E449F0C25 CRC64;

Query Match  
 Best Local Similarity 34.2%; Score 25; DB 5; Length 24;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 6 DERGMIP 13  
 DB 16 DVNGKIP 23

RESULT 12  
 ID O60696 PRELIMINARY: PRT: 20 AA.  
 AC O60696  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DE 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
 DE TYPE II AF10 PROTEIN (FRAGMENT).  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP Silliman C.C., McGavran L., Wei Q., Miller L.A., Li S., Hunger S.P.,  
 RL Leukemia 0:0-0(1998).  
 DR EMBL: AF060935; AAC16707.1; -.  
 FT NON\_TER 1 1 1  
 FT NON\_TER 1 1 1  
 SQ SEQUENCE 20 AA; 2188 MW; 34629142F7B03EA CRC64;

Query Match  
 Best Local Similarity 32.9%; Score 24; DB 4; Length 20;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 LIKDERG 9  
 DB 4 LVYCDHG 11

## RESULT 13

PRELIMINARY: PRT: 25 AA.  
 ID 077604  
 AC 077604  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DE TESTIS SPECIFIC PROTEIN (FRAGMENT).  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.  
 NCBI\_TaxID=9544;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP Harris E.E., Disotell T.R.;  
 RT "Nuclear Gene Trees and the Phylogenetic Relationships of the Mangabeys (Primates: Papionini)."  
 RL Mol. Biol. Evol. 0:0-0(1998).  
 DR EMBL: AF057416; AAC25144.1; -.  
 FT NON\_TER 1 1 1  
 FT NON\_TER 25 25 1  
 SQ SEQUENCE 25 AA; 2829 MW; A50DAC7A71719098 CRC64;

Query Match  
 Best Local Similarity 32.9%; Score 24; DB 6; Length 25;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 6 DERGMIPS 14  
 DB 5 DRGAIOS 13

RESULT 14  
 ID 077831 PRELIMINARY: PRT: 25 AA.  
 AC 077831  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DE 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE TESTIS SPECIFIC PROTEIN (FRAGMENT).  
 OS Cercopithecus mitis.  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Cercopithecus.  
 NCBI\_TaxID=36225;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-KIBBLES, YOYO;  
 RC Harris E.E., Disotell T.R.;  
 RT "Nuclear Gene Trees and the Phylogenetic Relationships of the Mangabeys (Primates: Papionini)."  
 RL Mol. Biol. Evol. 0:0-0(1998).  
 DR EMBL: AF057420; AAC25148.1; -.  
 DR EMBL: AF057419; AAC25147.1; -.  
 FT NON\_TER 1 1 1  
 FT NON\_TER 25 25 1  
 SQ SEQUENCE 25 AA; 2829 MW; A50DAC7A71719098 CRC64;

Query Match  
 Best Local Similarity 32.9%; Score 24; DB 6; Length 25;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 6 DERGMIPS 14  
 DB 5 DRGAIOS 13

## RESULT 15



```

077832
ID 077832 PRELIMINARY; PRT; 25 AA.
AC 077832;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE TESTIS SPECIFIC PROTEIN (FRAGMENT).
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Cercopithecus aethiops: Chordata: Craniata: Vertebrata: Euteleostomi;
OC Eukaryota; Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
OC Mammalia: Eutheria: Primates: Catarrhini: Cercopithecoidea;
OC Cercopithecoidea: Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=950004, 950003;
RA Harris E.E., Disotell T.R.;
RT "Nuclear Gene Trees and the Phylogenetic Relationships of the
RT Mangabey's (Primates: Papionini).";
RL Mol. Biol. Evol. 0:0-0(1998).
DR EMBL; AF057418; AAC25146.1; -.
DR EMBL; AF057417; AAC25145.1; -.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA: 2829 MW: A50DAC7A1719098 CRC64;

Query Match 32.9%; Score 24; DB 6; Length 25;
Best Local Similarity 55.6%; Pred. No. 1.6e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 DERGKMP 14
DB 5 DRRGATIG 13

```

```

RESULT 16
ID 09NL65 PRELIMINARY; PRT; 11 AA.
AC 09NL65;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ASAB-DELTA (FRAGMENT).
GN ASAB-DELTA.
OS Ascaris suum (pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa: Nematoda: Chromadorea; Ascaridoidea;
OC Ascaridoidea: Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE FROM N.A.
RA Kato Y.;
RT "Ascaris suum asabf-delta gene, exon 2."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029815; BAA89496.1; -.
FT NON_TER 1 1
FT NON_TER 11 11
SQ SEQUENCE 11 AA: 1187 MW: 8BADDCDIEAB5861 CRC64;

Query Match 31.5%; Score 23; DB 5; Length 11;
Best Local Similarity 60.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 CDERG 9
DB 7 CERKG 11

RESULT 17
ID 09TWF7 PRELIMINARY; PRT; 21 AA.
AC 09TWF7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE 27 KDA CYSTEINE PROTEINASE (FRAGMENT).
OS Entamoeba histolytica.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=5759;
RN [1]
RP SEQUENCE.
RX MEDLINE=96009779; PubMed=7558332;
RA Li E., Yang W.G., Zhang T., Stanley S.L., Jr.;
RT "Interaction of laminin with Entamoeba histolytica cysteine
RT proteinases and its effect on amebic pathogenesis."
RL Infect. Immun. 63:4150-4153(1995).
SQ SEQUENCE 21 AA: 2309 MW: 8C99CB65F1E7F834 CRC64;

```

```

Query Match 31.5%; Score 23; DB 5; Length 21;
Best Local Similarity 60.0%; Pred. No. 2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 6 DER--GKMP 13
DB 6 DERKGVTP 15

```

```

RESULT 18
ID 014347 PRELIMINARY; PRT; 23 AA.
AC 014347;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (FRAGMENT).
GN GAPD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85087928; PubMed=6096821;
RA Arcari P., Martinelli R., Salvatore F.;
RT "The complete sequence of a full length cDNA for human liver
RT glyceraldehyde-3-phosphate dehydrogenase: evidence for multiple mRNA
RT species."
RL Nucleic Acids Res. 12:9179-9189(1984).
DR EMBL; K03121; AAA52495.1; -.
DR HSSP; P56649; IDSS.
DR INTERPRO; IPR000173; -.
DR PFWA; PF00044; gpdm; 1.
FT NON_TER 23 23
SQ SEQUENCE 23 AA: 2311 MW: F413DDAF577717AA CRC64;

```

```

Query Match 31.5%; Score 23; DB 4; Length 23;
Best Local Similarity 80.0%; Pred. No. 2.2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 GKMP 13
DB 17 GKVP 21

```

```

RESULT 19
ID 09PRT7 PRELIMINARY; PRT; 23 AA.
AC 09PRT7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE MYOTOXIN IV (FRAGMENT).
OS Bochrops asper (Terciopelo).
OC Eukaryota; Metazoa: Chordata: Craniata: Vertebrata; Euteleostomi;

```

OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;  
 OC Viperidae; Crotalinae; Bothrops.  
 NCBI\_TaxID=8722;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95268655; PubMed=7749580;  
 RA Diaz C., Lomonte B., Zamudio F., Gutierrez J.M.;  
 RT "Purification and characterization of myotoxin IV, a phospholipase A2  
 variant, from Bothrops asper snake venom."  
 RL Nat. Toxins 3:26-31(1995).  
 DR HSP, P24605; ICLP.  
 SQ SEQUENCE 23 AA; 2512 MW; 8D84B07051C6D250 CRC64;

Query Match 31.5%; Score 23; DB 13; Length 23;  
 Best Local Similarity 83.3%; Pred. No. 2.2e+03;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ERGKM 12  
 Db 4 ELGKM 9

RESULT 20  
 Q9R573 PRELIMINARY; PRT; 24 AA.  
 AC Q9R573;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE HYDROXYLAMINE OXIDOREDUCTASE (FRAGMENT).  
 OS Nitrosomonas europaea.  
 CC Bacteria; Proteobacteria; beta subdivision;  
 CC Ammonia-oxidizing bacteria; Nitrosomonas.  
 NCBI\_TaxID=915;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93315429; PubMed=8325841;  
 RA Arciero D.M., Hooper A.B.;  
 RT "Hydroxylamine oxidoreductase from Nitrosomonas europaea is a multimer  
 of an octa-heme subunit."  
 RL J. Biol. Chem. 268:14645-14654(1993).  
 DR HSP, Q50925; IFCJ.  
 SQ SEQUENCE 24 AA; 2663 MW; 7EC6694C5DCD9D64 CRC64;

Query Match 31.5%; Score 23; DB 2; Length 24;  
 Best Local Similarity 57.1%; Pred. No. 2.3e+03;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 ERGKM 13  
 Db 16 DRGKATP 22

RESULT 21  
 Q94373 PRELIMINARY; PRT; 24 AA.  
 AC Q94373;  
 DT 01-FEB-1997 (TREMBLrel. 02, Created)  
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)  
 DE PPI-LIKE SER/THR PROTEIN PHOSPHATASE (FRAGMENT).  
 OS Caenorhabditis elegans.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 CC Rhabditidae; Petaloderinae; Caenorhabditis.  
 NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zele T., Gergely P., Dombradi V.;  
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; Z77734; CAB01293.1; -.  
 FT NON\_TER 1 1

FT NON\_TER 24 24  
 SQ SEQUENCE 24 AA; 2743 MW; 9C973B637D32483A CRC64;

Query Match 31.5%; Score 23; DB 5; Length 24;  
 Best Local Similarity 66.7%; Pred. No. 2.3e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILIKD 6  
 Db 8 ILMTCD 13

RESULT 22  
 061461 PRELIMINARY; PRT; 25 AA.  
 AC 061461;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE CYTOCHROME P-450B (PHENOBARBITAL-INDUCIBLE).  
 GN CYP2B13.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84207435; PubMed=6547088;  
 RA Stupans I., Ikeda T., Kessler D.J., Nebert D.W.;  
 RT "Characterization of a cDNA clone for mouse phenobarbital-inducible  
 cytochrome P-450b."  
 RL DNA 3:129-137(1984).  
 DR EMBL; K02409; AAA37510.1; -.  
 DR MGD; MGI:885899; CYP2B13.  
 DR INTERPRO; IPR001128; -.  
 DR PFM; PF00067; P430; 1.  
 DR NON\_TER 1 1  
 FT NON\_TER 25 25  
 SQ SEQUENCE 25 AA; 2979 MW; B9DF7A57355A1B0 CRC64;

Query Match 31.5%; Score 23; DB 11; Length 25;  
 Best Local Similarity 66.7%; Pred. No. 2.4e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 DERGM 11  
 Db 8 DDRKM 13

RESULT 23  
 Q9TRN8 PRELIMINARY; PRT; 15 AA.  
 AC Q9TRN8;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE 67 KDA ZONA PELLUCIDA BINDING PROTEIN HOMOLOG (FRAGMENT).  
 OS Sus scrofa (Pig).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92378826; PubMed=1510840;  
 RA Parry R.V., Barker P.F., Jones R.;  
 RT "Characterization of low Mr zona pellucida binding proteins from boar  
 spermatozoa and seminal plasma."  
 RL Mol. Reprod. Dev. 33:108-115(1992).  
 SQ SEQUENCE 15 AA; 1593 MW; 176B80F981EF1858 CRC64;

Query Match 30.1%; Score 22; DB 6; Length 15;  
 Best Local Similarity 57.1%; Pred. No. 2.2e+03;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 ERGKMP 13  
 |||:  
 Db 3 EAGPLIP 9

Db :|:  
 6 VLSRCD 11

Search completed: February 5, 2001, 11:27:26  
 Job time: 1879 sec

## RESULT 24

O9TRN3 PRELIMINARY; PRT; 15 AA.  
 AC O9TRN3;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, last annotation update)  
 DE LOW MR ZONA PELLUCIDA BINDING PROTEIN (FRAGMENT).  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92378826; PubMed=1510840;  
 RA Parry R.V., Barker P.J., Jones R.;  
 RT "Characterization of low Mr zona pellucida binding proteins from boar  
 spermatozoa and seminal plasma."  
 RL Mol. Reprod. Dev. 33:108-115(1992).  
 SQ SEQUENCE 15 AA; 1648 MM; 714A10FDACCA610E CRC64;

Query Match 30.1%; Score 22; DB 6; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 DERG 9  
 |||:  
 Db 7 DERG 10

## RESULT 25

O77922 PRELIMINARY; PRT; 16 AA.  
 AC O77922;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, last annotation update)  
 DE MHC CLASS II B LOCUS 16 (FRAGMENT).  
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).  
 OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;  
 OC Cichlidae; Oreochromis.  
 OX NCBI\_TaxID=8128;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98315113; PubMed=9649539;  
 RA Malaga-Trillo E., McAndrew B., Vincek V., Zaleska-Rutczynska Z.,  
 RA Suelmann H., Figueroa F., Klein J.;  
 RT "Linkage relationships and haplotype polymorphism among cichlid Mhc  
 class II B loci".  
 RL Genetics 149:1527-1547(1998).  
 DR EMBL; AF050035; AAC41374.1; -.  
 KW MHC.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 16 AA; 1893 MM; A56619E46B85F2D CRC64;

Query Match 30.1%; Score 22; DB 7; Length 16;  
 Best Local Similarity 50.0%; Pred. No. 2.3e+03;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ILIKCD 6



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 11:10:45 ; Search time 141.41 Seconds  
(without alignments)  
6.722 Million cell updates/sec

Title: US-09-343-406-19

Sequence: 73  
1 ILIKCDERGMIPS 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 5001

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: PIR\_66:\*\*

1: PIR1:\*\*  
2: PIR2:\*\*  
3: PIR3:\*\*  
4: PIR4:\*\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	31	42.5	16	2	PH1790	T cell receptor al
2	26	35.6	14	2	C44823	synaptosomal-assoc
3	26	35.6	20	2	A31516	lectin, galactose/
4	26	35.6	24	2	S68391	H+-transporting AT
5	25	34.2	13	2	S36668	hypothetical prote
6	25	34.2	17	2	S15778	insulin chain B-
7	24	32.9	19	2	A61110	68k collagen-bindl
8	23	31.5	16	2	S22677	heat shock protein
9	23	31.5	17	2	B20242	pyruvate kinase (E
10	23	31.5	24	2	T42258	Ser/Thr protein ph
11	23	30.8	25	2	A21630	cytochrome P450b-
12	22	30.1	20	2	S05411	carboxylesterase (
13	22	30.1	16	2	S00123	dihydrolipamide S
14	22	30.1	18	2	B24735	glutathione transf
15	22	30.1	20	2	PT0248	ig heavy chain CDR
16	21	28.8	14	2	PS0371	hypothetical prote
17	21	28.8	14	2	S33801	chaperone, TCP1-re
18	21	28.8	14	2	A17150	glucose 1-dehydrog
19	21	28.8	15	2	D48394	major fat-globule
20	21	28.8	20	1	LFBS7U	let leader peptide
21	21	28.8	20	2	S29100	glutathione transf
22	21	28.8	20	2	S23742	tet leader peptide
23	21	28.8	20	2	S65605	dimeric protein (B
24	21	28.8	21	2	A60975	HMGI/2 homolog - a
25	21	28.8	22	2	UT0581	natriuretic peptid
26	21	28.8	22	2	S11351	spore storage prot
27	20	27.4	10	2	PH0933	T-cell receptor be
28	20	27.4	12	2	S48209	zinc-metallophosph
29	20	27.4	12	2	S00706	actin - sea urchin

30	20	27.4	15	2	PH1788	T cell receptor al
31	20	27.4	16	2	A36889	leu operon leader
32	20	27.4	16	2	S65709	major allergen Myr
33	20	27.4	16	2	A44823	synaptosomal-assoc
34	20	27.4	16	2	PH1637	ig H chain V-D-J r
35	20	27.4	17	2	E59137	protein P103 - gol
36	20	27.4	18	2	B61110	68k collagen-bindl
37	20	27.4	18	2	S09026	carboxylesterase (
38	20	27.4	18	2	S18386	110k protein - mou
39	20	27.4	20	2	S30381	glutathione transf
40	20	27.4	20	2	A39328	notchlike tr-5b non
41	20	27.4	20	2	A60525	lysosome (EC 3.2.1
42	20	27.4	20	2	PC4387	GroEL protein homo
43	20	27.4	21	2	A60225	pyruvate dehydroge
44	20	27.4	22	1	MXKN1	mu-conotoxin GIIIA
45	20	27.4	22	1	MXKN2	mu-conotoxin GIIIB

## ALIGNMENTS

RESULT 1  
PH1790  
T cell receptor alpha chain V region (clone 2PBL V alpha 24-6) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Jul-1999  
C:Accession: PH1790  
R:Porcell, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.  
J. Exp. Med. 178, 1-16, 1993  
A:Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blo  
A:Reference number: PH1754; MUID:93301585  
A:Accession: PH1790  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-16 <POR>

Query Match 42.5%; Score 31; DB 2; Length 16;  
Best Local Similarity 50.0%; Pred. No. 48;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 CDERGMIPS 14  
DB 3 CDEGRIPPT 12

RESULT 2  
C44823  
synaptosomal-associated protein SNAP-25 peptide 8 - rabbit (fragment)  
N:Alternate names: superprotein peptide 8  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 31-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 15-Jun-1996  
C:Accession: C44823

R:Loewy, A.; Liu, W.S.; Baitinger, C.; Willard, M.B.  
J. Neurosci. 11, 3412-3421, 1991  
A:Title: The major 35S-methionine-labeled rapidly transported protein (superprotein)

A:Reference number: A44823; MUID:92044785  
A:Accession: C44823

A>Status: preliminary  
A:Molecule type: protein

A:Residues: 1-14 <LOE>  
A:Experimental source: visual tissue

A>Note: sequence extracted from NCBI backbone (NCBIP:64253)  
C:Keywords: membrane trafficking

Query Match 35.6%; Score 26; DB 2; Length 14;  
Best Local Similarity 36.4%; Pred. No. 3.1e+02;  
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 2 ILIKCDERGMIP 12  
|: |:|:|:

Db 2 LVMLDEGEQL 12

RESULT 3

A31516

lectin, galactose/N-acetylgalactosamine-specific - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 18-Jun-1993

C:Accession: A31516

R:Li, M.; Kawasaki, T.; Yamashina, I.

Biochem. Biophys. Res. Commun. 155, 720-725, 1988

A:Title: Structural similarity between the macrophage lectin specific for galactose/N-ac

A:Reference number: A31516; MUID:88339956

A:Accession: A31516

A:Molecule type: protein

A:Residues: 1-20 <IIX>

Query Match 35.6%; Score 26; DB 2; Length 20;

Best Local Similarity 60.0%; Pred. No. 4.3e+02;

Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 KDERGKMP 13

Db 8 KNDEAGKAP 17

RESULT 4

S68391

H+-transporting ATP synthase (EC 3.6.1.34) delta chain, chloroplast - Chlamydomonas rein

H:Alternate names: ATP synthase delta chain

C:Species: Chlamydomonas reinhardtii

C>Date: 04-Dec-1997 #sequence\_revision 12-Dec-1997 #text\_change 17-Mar-1999

C:Accession: S68391

R:Fieldner, H.R.; Schmid, R.; Leu, S.; Shavit, N.; Strommann, H.

FEBS Lett. 377, 163-166, 1995

A:Title: Isolation of CF(0)CF(1) from Chlamydomonas reinhardtii cw15 and the N-terminal

A:Reference number: S68388; MUID:96128220

A:Accession: S68391

A:Molecule type: protein

A:Residues: 1-24 <FIE>

A:Experimental source: strain CW15

C:Genetics:

A:Genome: nuclear

C:Superfamily: H+-transporting ATP synthase delta chain

C:Keywords: ATP biosynthesis; chloroplast; hydrolase; membrane-associated complex; thyla

Query Match 35.6%; Score 26; DB 2; Length 24;

Best Local Similarity 66.7%; Pred. No. 5.1e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 DERGKM 11

Db 18 DEGRKL 23

RESULT 5

S3668

hypothetical protein 133 - Plectonema boryanum (fragment)

C:Species: Plectonema boryanum

C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 15-Oct-1999

C:Accession: S3668

R:Fujita, Y.; Matsumoto, H.; Takahashi, Y.; Matsubara, H.

submitted to the EMBL Data Library, August 1992

A:Description: Identification of the nifDK-like gene (ORF467) involved in the biosynthes

A:Reference number: S36688

A:Accession: S36688

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-13 <FUV>

A:Cross-references: EMBL:D12973; NID:9216810; PIDN:BAA02348.1; PID:01002847; PID:9216811

Query Match 34.2%; Score 25; DB 2; Length 13;

Best Local Similarity 40.0%; Pred. No. 4.3e+02;

Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 CDERGKMPS 14

Db 2 CEGRRRTTPS 11

RESULT 6

S15778

Insulin chain B - bovine (fragments)

C:Species: Bos primigenius indicus x Bos primigenius taurus (cattle)

C>Date: 19-Mar-1997 #sequence\_revision 16-Jul-1999 #text\_change 16-Jul-1999

C:Accession: S15778; S15779

R:Bergman, T.; Agerberth, B.; Joernvall, H.

FEBS Lett. 283, 100-103, 1991

A:Title: Direct analysis of peptides and amino acids from capillary electrophoresis.

A:Reference number: S15778; MUID:91243852

A:Accession: S15778

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-8 <FEB1>

A:Accession: S15779

A:Status: preliminary

A:Molecule type: protein

A:Residues: 9-17 <FEB2>

C:Superfamily: Insulin

C:Keywords: hormone; pancreas

Query Match 34.2%; Score 25; DB 2; Length 17;

Best Local Similarity 44.4%; Pred. No. 5.5e+02;

Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILIKDEG 9

Db 2 LYLWCGEG 10

RESULT 7

A61110

68k collagen-binding protein, heavy form - chicken (fragments)

C:Species: Gallus gallus (Chicken)

C>Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-Oct-1994

C:Accession: A61110

R:Tidball, J.G.

J. Biol. Chem. 267, 21211-21219, 1992

A:Title: Identification and distribution of a novel, collagen-binding protein in the

A:Reference number: A61110; MUID:93016046

A:Accession: A61110

A:Molecule type: protein

A:Residues: 1-19 <TID>

C:Keywords: collagen binding

Query Match 32.9%; Score 24; DB 2; Length 19;

Best Local Similarity 40.0%; Pred. No. 9.1e+02;

Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 IKDERGKMI 12

Db 2 VTLDSEGRKV 11

RESULT 8

S22677

heat shock protein groEL - Pseudomonas maltophilia (fragment)

C:Species: Pseudomonas maltophilia

C>Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 17-Mar-1999

C:Accession: S22677

R:Fowell, S.L.; Lilley, K.S.; Jones, D.; Maxwell, A.

Mol. Microbiol. 6, 1575-1576, 1992  
A:Title: GroEL proteins from three *Pseudomonas* species.  
A:Reference number: S22676; MUID:92326643  
A:Accession: S22677  
A:Molecule type: protein  
A:Residues: 1-16 <FW>  
C:Genetics:  
A:Gene: groEL  
C:Keywords: molecular chaperone

Query Match 31.5%; Score 23; DB 2; Length 16;  
Best Local Similarity 57.1%; Pred. No. 1.2e+03;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 6 DERGKM 12  
DB 10 DARSKM 16

RESULT 9  
B20242  
pyruvate kinase (EC 2.7.1.40) - rainbow trout (tentative sequence) (fragment)  
C:Species: *Oncorhynchus mykiss* (rainbow trout)  
C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 31-Mar-2000  
C:Accession: B20242  
R:Harkins, R.N.; Noccon, J.C.; Russell, M.P.; Fothergill, L.A.; Muirhead, H.  
Eur. J. Biochem. 136, 341-346, 1983  
A:Title: A comparison of the structure and activity of cat and trout muscle pyruvate kin  
A:Reference number: A94659; MUID:84028655  
A:Accession: B20242  
A:Molecule type: protein  
A:Residues: 1-17 <HAR>  
C:Function:  
A:Description: catalyzes the transphosphorylation of phosphoenolpyruvate and ADP to pyr  
A:Pathway: glycolysis  
C:Superfamily: pyruvate kinase  
C:Keywords: ATP biosynthesis; glycolysis; magnesium; metalloprotein; phosphotransferase;

Query Match 31.5%; Score 23; DB 2; Length 17;  
Best Local Similarity 33.3%; Pred. No. 1.2e+03;  
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 KCDERGKM 12  
DB 3 RCRNRAKPV 11

RESULT 10  
T42258  
Ser/Thr protein phosphatase homolog - *Caenorhabditis elegans* (fragment)  
C:Species: *Caenorhabditis elegans*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 02-Jun-2000  
C:Accession: T42258  
R:Zeke, T.; Gergely, P.; Dombradi, V.  
submitted to the EMBL Data Library, July 1996  
A:Description: The catalytic subunits of Ser/Thr protein phosphatases from *Caenorhabditis*  
A:Reference number: 222131  
A:Accession: T42258  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-24 <ZEK>  
A:Cross-references: EMBL:T77734; PIDN:CA801293.1  
C:Superfamily: phosphoprotein phosphatase; phosphoesterase core homology; phosphoprotein

Query Match 31.5%; Score 23; DB 2; Length 24;  
Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ILIKCD 6  
DB 11

DB 8 ILMTCD 13

RESULT 11  
A21630  
cytochrome P450b - mouse (fragment)  
N:Contains: oxidoreductase (EC 1.-.-.-)  
C:Species: *Mus musculus* (house mouse)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 11-Jun-1999  
C:Accession: A21630  
R:Stupans, I.; Ikeda, T.; Kessler, D.J.; Nebert, D.W.  
DNA 3, 129-137, 1984  
A:Title: Characterization of a cDNA clone for mouse phenobarbital-inducible cytochrom  
A:Reference number: A21630; MUID:84207435  
A:Accession: A21630  
A:Molecule type: mRNA  
A:Residues: 1-25 <STD>  
A:Cross-references: GB:K02409; NID:g192896; PIDN:AAA37510.1; PID:g553905  
A:Note: the authors translated the codon CTA for residue 5 as Pro, ACC for residue 7  
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
C:Keywords: electron transfer; heme; monooxygenase; oxidoreductase; transmembrane pro

Query Match 31.5%; Score 23; DB 2; Length 25;  
Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 DERGKM 11  
DB 8 DDRSKM 13

RESULT 12  
S05411  
carboxylesterase (EC 3.1.1.1) - *Sulfolobus acidocaldarius* (fragment)  
N:Alternate names: serine esterase  
C:Species: *Sulfolobus acidocaldarius*  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 28-Apr-1993  
C:Accession: S05411  
R:Sober, H.; Goerlich, H.  
Biochem. J. 261, 993-998, 1989  
A:Title: Further kinetic and molecular characterization of an extremely heat-stable c  
A:Reference number: S05411; MUID:90026296  
A:Accession: S05411  
A:Molecule type: protein  
A:Residues: 1-20 <SOB>  
A:Note: 1-Ala and 1-Ser were also found  
C:Keywords: carboxylic ester hydrolase; tetramer

Query Match 30.8%; Score 22.5; DB 2; Length 20;  
Best Local Similarity 58.3%; Pred. No. 1.7e+03;  
Matches 7; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

OY 3 IKC-DEGKMP 13  
DB 6 IKCLSGFVIP 17

RESULT 13  
S00123  
dihydrolypamide S-succinyltransferase (EC 2.3.1.61) - bovine (fragment)  
N:Alternate names: 2-oxoglutarate dehydrogenase complex chain E2; transsuccinylase  
C:Species: *Bos primigenius taurus* (cattle)  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 05-May-2000  
C:Accession: S00123  
R:Bradford, A.P.; Aitken, A.; Beg, F.; Cook, K.G.; Yeaman, S.J.  
FEBS Lett. 222, 211-214, 1987  
A:Title: Amino acid sequence surrounding the lipoic acid cofactor of bovine kidney 2-  
A:Reference number: S00123; MUID:88005143  
A:Accession: S00123  
A:Molecule type: protein  
A:Residues: 1-16 <BRA>

C:Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology  
C:Keywords: acyltransferase; coenzyme A; lipoamide; mitochondrion; oxidoreductase; triac  
F:1-16/Domain: lipoyl/biotin-binding homology (fragment) <PB>  
F:5/Binding site: lipoamide (Lys) (covalent) #status experimental

Query Match 30.1%; Score 22; DB 2; Length 16;  
Best Local Similarity 33.3%; Pred. No. 1.7e+03;  
Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ICDEKGMIPS 14  
| : : : :  
DB 1 IEDKTSVQVPS 12

RESULT 14  
B24735  
glutathione transferase (EC 2.5.1.18) 1-2 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 25-Oct-1987 #sequence\_revision 25-Oct-1987 #text\_change 18-Jun-1993  
C:Accession: B24735

R.Mannervik, B.; Alin, P.; Guthenberg, C.; Jansson, H.; Tahlir, M.K.; Warholm, M.; Jornvall  
Proc. Natl. Acad. Sci. U.S.A. 82: 7202-7206, 1985  
A:Title: Identification of three classes of cytosolic glutathione transferase common to  
A:Reference number: A24735; MUID:86042634  
A:Accession: B24735  
A:Molecule type: protein  
A:Residues: 1-18 <MAN>  
C:Superfamily: glutathione transferase  
C:Keywords: transferase

Query Match 30.1%; Score 22; DB 2; Length 18;  
Best Local Similarity 66.7%; Pred. No. 1.9e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 RGMIP 13  
| : : : :  
DB 12 RGMIP 17

RESULT 15  
PT0248  
Ig heavy chain CDR3 region (clone 2-106B) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 06-Jun-1997  
C:Accession: PT0248  
R.Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and  
A:Reference number: PT0222; MUID:91108337  
A:Accession: PT0248  
A:Molecule type: DNA  
A:Residues: 1-20 <YAM>  
A:Experimental source: B lymphocyte  
C:Superfamily: immunoglobulin V region: immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin

Query Match 30.1%; Score 22; DB 2; Length 20;  
Best Local Similarity 60.0%; Pred. No. 2.1e+03;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 CDERG 9  
| : : : :  
DB 1 CDRG 5

RESULT 16  
PS0371  
hypothetical protein (psac region) - Synechococcus sp. (fragment)  
C:Species: Synechococcus sp.  
C:Date: 17-Aug-1992 #sequence\_revision 17-Aug-1992 #text\_change 08-Oct-1999

C:Accession: PS0371  
R.Rhiel, E.; Stirewalt, V.L.; Gasparich, G.E.; Bryant, D.A.  
Gene 112, 123-128, 1992  
A:Title: The psac genes of Synechococcus sp. PCC7002 and Cyanophora paradoxa: cloning  
A:Reference number: JS0694; MUID:92201692  
A:Accession: PS0371  
A:Molecule type: DNA  
A:Residues: 1-14 <RHI>  
A:Cross-references: GB:M6238; NID:9154574; PIDN:AAA27351.1; PID:9552030

Query Match 28.8%; Score 21; DB 2; Length 14;  
Best Local Similarity 40.0%; Pred. No. 2.3e+03;  
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 KCDERGMIP 13  
| : : : :  
DB 5 KIDVTGRLOP 14

RESULT 17  
S33801  
chaperone, TCP1-related - oat  
C:Species: Avena sativa (oat)  
C:Date: 02-Dec-1993 #sequence\_revision 27-Feb-1997 #text\_change 17-Mar-1999  
C:Accession: S33801  
R.Nummert, E.; Grimm, R.; Speith, V.; Eckerskorn, C.; Schillitz, E.; Gatenby, A.A.; Scha  
Nature 363, 644-648, 1993  
A:Title: A TCP1-related molecular chaperone from plants refolds phycochrome to its ph  
A:Reference number: S33800; MUID:93288140  
A:Accession: S33801  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-14 <NUM>

Query Match 28.8%; Score 21; DB 2; Length 14;  
Best Local Similarity 42.9%; Pred. No. 2.3e+03;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 ERGMIP 13  
| : : : :  
DB 3 ESKLVP 9

RESULT 18  
A17150  
glucose 1-dehydrogenase (NAD+) (EC 1.1.1.118) - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 07-Feb-1997  
C:Accession: A17150  
R.Franzen, B.; Carubba, C.; Felngold, D.S.; Ashcom, J.; Franzen, J.S.  
Biochem. J. 199, 599-602, 1981  
A:Title: Amino acid sequence of the tryptic peptide containing the catalytic-site th  
A:Reference number: A17150; MUID:82182061  
A:Accession: A17150  
A:Molecule type: protein  
A:Residues: 1-14 <FRA>  
C:Keywords: NAD; oxidoreductase

Query Match 28.8%; Score 21; DB 2; Length 14;  
Best Local Similarity 50.0%; Pred. No. 2.3e+03;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 CDERG 10  
| : : : :  
DB 9 CFZGK 14

RESULT 19  
D8394  
major fat-globule membrane protein GP 55 - guinea pig (fragment)



C:Species: Cavia porcellus (guinea pig)  
C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 31-Oct-1997  
C:Accession: D48394  
R:Author: I.H.; Banghart, L.R.; Lane, W.S.  
Biochem. Mol. Biol. Int. 29, 545-554, 1993  
A:Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig  
II-like sequences.  
A:Reference number: A48394; MUID:93250576  
A:Accession: D48394  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-15 <MAT>  
A:Experimental source: milk  
A>Note: sequence extracted from NCBI Backbone (NCBI:131448)  
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo

Query Match 28.8%; Score 21; DB 2; Length 15;  
Best Local Similarity 33.3%; Pred. No. 2.4e+03;  
Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 ILICDERGKMP 13  
|||  
DB 3 ILGCEINGCLEP 14

RESULT 20  
LEBSTU  
tet leader peptide - Bacillus cereus plasmid pBC16  
C:Species: Bacillus cereus  
C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 16-Jul-1999  
C:Accession: S09233  
R:Palva, A.; Vidgren, G.; Simonen, M.; Rintala, H.; Laamanen, P.  
Nucleic Acids Res. 18, 1635, 1990  
A:Title: Nucleotide sequence of the tetracycline resistance gene of pBC16 from Bacillus  
A:Reference number: S09233; MUID:90221899  
A:Accession: S09233  
A>Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-20 <PAL>  
A:Cross-references: EMBL:X51366; NID:939459; PIDN:CAA35750.1; PID:939460  
C:Genetics:  
A:Genome: plasmid  
C:Superfamily: tet leader peptide

Query Match 28.8%; Score 21; DB 1; Length 20;  
Best Local Similarity 60.0%; Pred. No. 3.1e+03;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 IKCDE 7  
|||  
DB 1 MKCNE 5

RESULT 21  
S29100  
glutathione transferase (EC 2.5.1.18) class mu M2 - golden hamster (fragment)  
C:Species: Mesocricetus auratus (golden hamster)  
C:Date: 25-Feb-1994 #sequence\_revision 08-Nov-1996 #text\_change 07-May-1999  
C:Accession: S29100  
R:Bogaards, J.J.P.; van Ommen, B.; van Bladeren, P.J.  
Biochem. J. 286, 383-388, 1992  
A:Title: Purification and characterization of eight glutathione S-transferase isoenzymes  
A:Reference number: S29097; MUID:92412006  
A:Accession: S29100  
A:Molecule type: protein  
A:Residues: 1-20 <BOG>  
C:Comment: At least five species-independent classes of cytosolic glutathion transferas  
es mitochondrial form are known.  
C:Complex: dimer  
A:Function:  
A:Description: catalyzes the nucleophilic conjugation of intracellular glutathione to a

A:Pathway: detoxification; xenobiotics metabolism  
A>Note: Increased hydrophilicity of GSH-conjugates facilitates their further metaboli  
es of damage  
C:Superfamily: glutathione transferase  
C:Keywords: dimer; transferase

Query Match 28.8%; Score 21; DB 2; Length 20;  
Best Local Similarity 46.2%; Pred. No. 3.1e+03;  
Matches 6; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 ILICDERGKMP 13  
|||  
DB 3 ILXYWDVRLTYP 15

RESULT 22  
S23742  
tet leader peptide - Staphylococcus hyicus  
C:Species: Staphylococcus hyicus  
C:Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 26-Aug-1999  
C:Accession: S23742  
R:Schwarz, S.; Cardoso, M.; Wegener, H.C.  
Antimicrob. Agents Chemother. 36, 580-588, 1992  
A:Title: Nucleotide sequence and phylogeny of the tet(L) tetracycline resistance dete  
A:Reference number: S23742; MUID:92321725  
A:Accession: S23742  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-20 <SCH>  
A:Cross-references: EMBL:X60828; NID:943682; PIDN:CAA3219.1; PID:943683  
A>Note: the authors translated the codon GGA for residue 14 as Ala  
C:Superfamily: tet leader peptide

Query Match 28.8%; Score 21; DB 2; Length 20;  
Best Local Similarity 60.0%; Pred. No. 3.1e+03;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 IKCDE 7  
|||  
DB 1 MKCNE 5

RESULT 23  
S65605  
dimeric protein (BDP) - barley (fragment)  
C:Species: Hordeum vulgare (barley)  
C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 25-Apr-1997  
C:Accession: S65605  
R:Garcia-Casado, G.; Armentia, A.; Sanchez-Monge, R.; Sanchez, L.M.; Lopez-Otin, C.;  
FEBS Lett. 364, 36-40, 1995  
A:Title: A major baker's asthma allergen from rye flour is considerably more active t  
A:Reference number: S65604; MUID:95269763  
A:Accession: S65605  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-20 <GAR>

Query Match 28.8%; Score 21; DB 2; Length 20;  
Best Local Similarity 80.0%; Pred. No. 3.1e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 9 GKMP 13  
|||  
DB 11 GKSP 15

RESULT 24  
A60975  
HMGI/2 homolog - axolotl (fragment)  
C:Species: Ambystoma mexicanum (axolotl)

C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 31-Dec-1993  
C:Accession: A60975  
R:Guillet, F.; Tournefier, A.; Denoulet, P.; Capony, J.P.; Kerfourn, F.; Charlemagne, J.  
Biol. Cell 69, 153-160, 1990  
A:Title: High levels of HMG-2 protein expression in the cytoplasm and nucleus of hydroc  
A:Reference number: A60975; MUID:91265891  
A:Accession: A60975  
A:Molecule type: protein  
A:Residues: 1-21 <GUI>  
A>Note: this protein was found in the cytoplasmic and nuclear compartments of thymocytes  
C:Superfamily: nonhistone chromosomal protein HMG-2; HMG box homology  
F:5-21/Domain: HMG box homology (fragment) <HMG>

Query Match 28.8%; Score 21; DB 2; Length 22;  
Best Local Similarity 100.0%; Pred. No. 3.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 RGKM 11  
|||  
Db 9 RGKM 12

## RESULT 25

JT0581  
natriuretic peptide type C - chicken  
C:Species: Gallus gallus (chicken)  
C>Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 20-Mar-1998  
C:Accession: JT0581  
R:Arimura, J.J.; Minamino, N.; Kangawa, K.; Matsuo, H.  
Biochem. Biophys. Res. Commun. 174, 142-148, 1991  
A:Title: Isolation and identification of C-type natriuretic peptide in chicken brain.  
A:Reference number: JT0581; MUID:9113186  
A:Accession: JT0581  
A:Molecule type: protein  
A:Residues: 1-22 <ARI>  
A:Experimental source: brain  
C:Superfamily: natriuretic peptide A precursor  
C:Keywords: diuretic; hormone; natriuretic; osmoregulation

Query Match 28.8%; Score 21; DB 2; Length 22;  
Best Local Similarity 44.4%; Pred. No. 3.4e+03;  
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 IKCDERGM 11  
:| | | |  
Db 9 VKLDRI GSM 17

Search completed: February 5, 2001, 11:10:46  
Job time: 1057 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 11:34:15 ; Search time 78.94 Seconds  
(Without alignments)  
5.727 Million cell updates/sec

Title: US-09-343-406-19  
Perfect score: 73  
Sequence: 1 ILIKCDERGMIPS 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 1436

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	23	31.5	25 1 PA22_BOTSC	P80963 bothriechis
2	22	30.1	16 1 ODO2_BOVIN	P11179 bos taurus
3	21	28.8	19 1 FIBB_TARTF	P14539 tapirus ter
4	21	28.8	20 1 LPTB_BACST	P05658 bacillus st
5	21	28.8	22 1 ANFC_CHICK	P21805 gallus gall
6	20	27.4	8 1 ACT_CARMA	P80709 caracaus ma
7	20	27.4	17 1 RS13_PARDE	P72180 paracoccus
8	20	27.4	20 1 LYC_FELCA	P37155 felis silve
9	20	27.4	21 1 FIBB_ANTAM	P14465 antilocapra
10	20	27.4	21 1 FIBB_MUNMU	P14475 muntiacus m
11	20	27.4	21 1 FIBB_ODOHE	P14476 odocoiileus
12	20	27.4	21 1 FIBB_RANTA	P14479 rangifer ta
13	20	27.4	22 1 CXM1_CONGE	P01523 conus geogr
14	20	27.4	22 1 CXM2_CONGE	P01524 conus geogr
15	20	27.4	22 1 CXM3_CONGE	P05482 conus geogr
16	19.5	26.7	21 1 THAN_PODMA	P55788 podisus mac
17	19	26.0	16 1 BAITI_EUBSP	P32371 eubacterium
18	19	26.0	16 1 PGK_VIBCH	P96154 vibrio chol
19	19	26.0	17 1 ITHB_HIRME	P28502 hirtudo medi
20	19	26.0	20 1 LPTB_BACSU	P23053 bacillus su
21	19	26.0	20 1 PGK_CLOPA	P81346 clostridium
22	19	26.0	22 1 ACE_BOVIN	P12820 bos taurus
23	19	26.0	22 1 CH10_RHOSH	P25969 rhodobacter
24	19	26.0	22 1 CYTB_THETS	P81064 thermozon
25	18	24.7	15 1 UC08_MAIZE	P80614 zea mays (m
26	18	24.7	17 1 TRP2_LEUMA	P81733 leucophaea
27	18	24.7	19 1 SCX6_TITBA	P56610 titus bahl
28	18	24.7	20 1 FIBB_SHEEP	P14470 ovis aries
29	18	24.7	20 1 PYRR_PVRAP	P37362 pyrrhocoris
30	18	24.7	20 1 SUCB_CANFA	P99507 canis famli
31	18	24.7	21 1 FIBB_BOBBU	P14467 bubalus bub
32	18	24.7	23 1 UDP_LACCA	P19662 lactobacilli
33	18	24.7	25 1 ANDT_ANDAU	P56684 androctonus

34	17.5	24.0	24 1 FIBB_CANFA	P12800 canis famli
35	17	23.3	10 1 SLAP_BACTG	P49325 bacillus th
36	17	23.3	15 1 UC20_MAIZE	P80626 zea mays (m
37	17	23.3	16 1 ALRX_PSEPU	P17916 pseudomonas
38	17	23.3	17 1 CHH3_BOOMO	P20729 bombyx mori
39	17	23.3	17 1 SRY_URSAR	P36396 ursus arcto
40	17	23.3	19 1 FIBB_PIG	P14477 sus scrofa
41	17	23.3	19 1 LCRP_PETMA	P10996 petromyzon
42	17	23.3	20 1 UN05_PIPNS	P81674 pinus pins
43	17	23.3	21 1 FEDB_AMEYE	P80706 amycolatops
44	17	23.3	21 1 FIBB_BISBO	P14466 bison bonas
45	17	23.3	21 1 GIRA_STRSH	P50071 streptomyce

## ALIGNMENTS

```

RESULT 1
ID PA22_BOTSC STANDARD; PRT; 25 AA.
AC P80963;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PHOSPHOLIPASE A2 HOMOLOG, MYOTOXIN II (FRAGMENT).
OS Bothriechis schlegelii (Eyelash viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Scleroglossa; Serpentes; Colubridae;
OC Viperidae; Crotalinae; Bothriechis.
RN [1]
RP SEQUENCE.
RC TISSUE=VENOM;
RX MEDLINE=97209510; PubMed=9056257;
RA Angulo Y., Chaves E., Alape A., Rucavado A., Gutierrez J.M.,
RA Lomonte B.;
RT "Isolation and characterization of a myotoxic phospholipase A2 from
RT the venom of the arboreal snake Bothriechis (Bothrops) schlegelii
RT from Costa Rica."
RL Arch. Biochem. Biophys. 339:260-266(1997).
CC -!- FUNCTION: THIS MYOTOXIC PROTEIN LACKS ENZYMATIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
DR HSSP; P24605; ICLP.
DR INTERPRO: IPR001211; -.
DR FRAM; PF00068; Phos1lp.1.
DR PROSITE; PS00118; PA2_HIS; PARTIAL.
DR PROSITE; PS00119; PA2_ASP; PARTIAL.
KW Toxin; Venom; Multigene family.
FT NON_TER 25
SQ SEQUENCE 25 AA; 2725 MM; C578AB09235EA02A CRC64;

```

Query Match 31.5%; Score 23; DB 1; Length 25;  
Best Local Similarity 83.3%; Pred. No. 6.5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ERGKMI 12  
| | | | |  
Db 4 ELGKMI 9

```

RESULT 2
ID ODO2_BOVIN STANDARD; PRT; 16 AA.
AC P11179;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DIHYDROLIPOAMIDE SUCCINYLTRANSFERASE COMPONENT OF 2-OXOGLOUTARATE
DE DEHYDROGENASE COMPLEX (EC 2.3.1.61) (E2) (FRAGMENT).
GN D1ST.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

```

OC Bovidae; Bovinae; Bos.  
 RN [1]  
 RP SEQUENCE.  
 RX TISSUE=KIDNEY;  
 RA MEDLINE=88005143; PubMed=3115829;  
 RA Bradford A.P., Altken A., Beg F., Cook K.G., Yeaman S.J.;  
 RT "Amino acid sequence surrounding the lipolic acid cofactor of bovine  
 RT kidney 2-oxoglutarate dehydrogenase complex.";  
 RT FEBS Lett. 223:211-214(1987)  
 CC -1- FUNCTION: THE 2-OXOGLUTARATE DEHYDROGENASE COMPLEX CATALYZES THE  
 CC OVERALL CONVERSION OF 2-OXOGLUTARATE TO SUCCINYL-COA & CO(2). IT  
 CC CONTAINS MULTIPLE COPIES OF 3 ENZYMAIC COMPONENTS: 2-OXOGLUTARATE  
 CC DEHYDROGENASE (E1), DIHYDROLIPOAMIDE SUCCINYLTRANSFERASE (E2) &  
 CC LIPOAMIDE DEHYDROGENASE (E3).  
 CC -1- CATALYTIC ACTIVITY: SUCCINYL-COA + DIHYDROLIPOAMIDE = COA +  
 CC S-SUCCINYLDIHYDROLIPOAMIDE.  
 CC -1- COFACTOR: THE E2 COMPONENT CONTAINS ONE COVALENTLY-BOUND LIPOYL  
 CC COFACTOR.  
 CC -1- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL  
 CC SYMMETRY.  
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.  
 CC -1- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 LIPOYL-BINDING DOMAIN.  
 CC PIR: S00123; S00123.  
 DR INTERPRO: IPR003016; -.  
 DR PROSITE: PS00189; LIPOYL, PARTIAL.  
 KM GLYCOLYSIS; Transferase; Acyltransferase; Lipoyl; Mitochondrion.  
 FT NON\_TER 1 5 1  
 FT BINDING 1 5 5  
 FT SEQUENCE 16 AA: 1643 MW: E34CD29B86B3239D CRC64;  
 SQ  
 Query Match 30.1%; Score 22; DB 1; Length 16;  
 Best Local Similarity 33.3%; Pred. No. 6.3e+02;  
 Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 3 INCDEKMIIPS 14  
 I: I: I: I:  
 Db 1 IETDKTSVQVPS 12

RESULT 3  
 FIBB\_TAPPE STANDARD; PRT; 19 AA.  
 ID FIBB\_TAPPE  
 AC P14539;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE FIBRINOPEPTIDE B.  
 OS Tapirus terrestris (Lowland tapir) (Brazilian tapir).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Tapiridae; Tapirus.  
 RN [1]  
 RA SEQUENCE.  
 RP O'Neill P.B., Doolittle R.F.;  
 RT "Mammalian phylogeny based on fibrinopeptide amino acid sequences.";  
 RT Syst. Zool. 22:550-595(1973).  
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA, & GAMMA). LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 DR INTERPRO: IPR002181; -.  
 DR PROSITE: PS00514; FIBRIN, AG.C DOMAIN; PARTIAL.  
 KM Blood coagulation; Plasma; Sulfatation.  
 FT MOD\_RES 4 4  
 FT NON\_TER 19 19  
 FT SEQUENCE 19 AA: 2226 MW: FCEA265G35E99627 CRC64;  
 SQ

Query Match 28.8%; Score 21; DB 1; Length 19;  
 Best Local Similarity 80.0%; Pred. No. 1.1e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 DERK 10  
 I: I: I: I:  
 Db 9 DERAK 13

RESULT 4  
 LPTR\_BACST STANDARD; PRT; 20 AA.  
 ID LPTR\_BACST  
 AC P05658;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 01-FEB-1995 (Rel. 31, Last annotation update)  
 DE TETRACYCLINE RESISTANCE LEADER PEPTIDE.  
 GN TETL.  
 OS Bacillus stearothermophilus, Bacillus cereus, and  
 OS Staphylococcus hyicus  
 OC Plasmid pPH15, Plasmid pBC16, and Plasmid pSTEL.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC PLASMID-PH15;  
 RX MEDLINE=86031344; PubMed=2969683;  
 RA Hoshino T., Ikeda T., Tomizuka N., Furukawa K.;  
 RT "Nucleotide sequence of the tetracycline resistance gene of pPH15, a  
 RT thermophilic Bacillus plasmid: comparison with staphylococcal TCR  
 RT controls.";  
 RT Gene 37:131-138(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC PLASMID-PBC16;  
 RX MEDLINE=90221899; PubMed=2109312;  
 RA Palva A., Vidgren G., Simonen M., Rintala H., Laamanen P.;  
 RT "Nucleotide sequence of the tetracycline resistance gene of pBC16  
 RT from Bacillus cereus.";  
 RT Nucleic Acids Res. 18:1635-1635(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-S\_HYICUS; PLASMID-PSTEL;  
 RX MEDLINE=92321725; PubMed=1622166;  
 RA Schwarz S., Cardoso M., Wegener H.C.;  
 RT "Nucleotide sequence and phylogeny of the tet(L) tetracycline  
 RT resistance determinant encoded by plasmid pSTEL from Staphylococcus  
 RT hyicus.";  
 RT Antimicrob. Agents Chemother. 36:580-588(1992).  
 CC -----  
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; D00006; BAA00004.1; -;  
 DR EMBL; M1036; AAA22850.1; -;  
 DR EMBL; X51366; CAA35750.1; -;  
 DR EMBL; X60828; CAA43219.1; -;  
 DR PIR: S09233; LEFSTV.  
 DR PIR: S23742; S23742.  
 KM Leader peptide; Antibiotic resistance; Plasmid.  
 SQ SEQUENCE 20 AA: 2253 MW: 18D0F46CA23JCA1 CRC64;  
 SQ

Query Match 28.8%; Score 21; DB 1; Length 20;  
 Best Local Similarity 60.0%; Pred. No. 1.2e+03;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 IKCDE 7  
:1:1  
DB 1 MKCNE 5

## RESULT 5

ANFC\_CHICK STANDARD; PRT: 22 AA.

AC P21805; 01-MAY-1991 (Rel. 18, Created)

DT 01-MAY-1991 (Rel. 18, Last sequence update)

DE C-TYPE NATRIURETIC PEPTIDE (CNP).

GN NPCC.

OS Gallus gallus (chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

CC [1]

RP SEQUENCE.

RC TISSUE-BRAIN:

RA MEDLINE-91113186; PubMed-1989595.

RT Altmura J.J., Minamino N., Kangawa K., Matsuo H.;

"Isolation and identification of C-type natriuretic peptide in chicken brain."

RL Biochem. Biophys. Res. Commun. 174:142-148(1991).

CC -1- FUNCTION: VASORELAXANT ACTIVITY. HAS A CGMP-STIMULATING ACTIVITY.

CC -1- SIMILARITY: BELONGS TO THE NATRIURETIC PEPTIDES FAMILY.

DR PIR: J0581, J0581.

DR INTERPRO: IPR000663; -.

DR PFAM: PF00212; ANP; 1.

DR PRINTS: PR00710; NATPEPTIDES.

DR PROSITE: PS00263; NATRIURETIC\_PEPTIDE; 1.

KW Vasoactive.

FT DISULFID

FT SEQUENCE 22 AA; 2244 MW; FC2A4706DBDAC025 CRC64;

SO

Query Match 28.8%; Score 21; DB 1; Length 22;

Best Local Similarity 44.4%; Pred. No. 1.3e+03;

Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 IKCDERGM 11

DB 9 VKLDRIISM 17

RESULT 6

ACT\_CARMA STANDARD; PRT: 8 AA.

AC P80709; 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DE 15-JUL-1999 (Rel. 38, Last annotation update)

DE ACTIN (FRAGMENT).

OS Carcarius maenas (Common shore crab) (Green crab).

OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;

OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;

OC Eubrachyura; Portunoidae; Portunidae; Carcinus.

CC [1]

RP SEQUENCE.

RA Lachaise F., Somme G., Carpentier G., Granjeon E., Webster S.;

BAchdassarian D.;

RT "A transaldolase, An enzyme implicated in crab steroidogenesis.";

RL Endocrine 5:23-32(1996).

CC -1- FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED

CC IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBQUITOUSLY EXPRESSED

CC IN ALL EUKARYOTIC CELLS.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:

CC 6.8, ITS MW IS: 46 KDA.

CC -1- SIMILARITY: BELONGS TO THE ACTIN FAMILY.

DR INTERPRO: IPR000279; -

DR PROSITE: PS00406; ACTINS\_1; PARTIAL.

DR PROSITE: PS00432; ACTINS\_2; PARTIAL.

DR PROSITE: PS01132; ACTINS\_ACT\_LIKE; PARTIAL.

KW Structural protein.

FT NON\_TER 1 1

FT NON\_TER 8 8

FT SEQUENCE 8 AA; 976 MW; 1424005AB2CAAB3 CRC64;

SO

Query Match 27.4%; Score 20; DB 1; Length 8;

Best Local Similarity 100.0%; Pred. No. 8.9e+04;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KCD 6

DB 1 KCD 3

RESULT 7

RS13\_PARDE STANDARD; PRT: 17 AA.

AC P72180; 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DE 30-MAY-2000 (Rel. 39, Last annotation update)

DE 30S RIBOSOMAL PROTEIN S13 (FRAGMENT).

GN RPSM.

OS Paracoccus denitrificans.

OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum rubrum group;

OC Paracoccus.

CC [1]

RP SEQUENCE FROM N.A.

RC STRAIN-71.11F.

RA Perrier V., Bouscass A., Meier O., Barzu O., Gilles A.M.;

"Adenylate kinase from P. denitrificans: an iron and zinc binding

RT protein, catalyzes phosphorylation of AMP and reduction of

RT cytochrome C."

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: INVOLVED IN THE BINDING OF FMET-TRNA AND, HENCE, IN THE

CC INITIATION OF TRANSLATION (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE S13P FAMILY OF RIBOSOMAL PROTEINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: U64204; AAB06329.1; ALT\_INIT.

DR INTERPRO: IPR001892; -.

DR PROSITE: PS00646; RIBOSOMAL\_S13; PARTIAL.

KW Ribosomal protein.

FT NON\_TER 17 17

FT SEQUENCE 17 AA; 1793 MW; 43F7F60DC8670B7D CRC64;

SO

Query Match 27.4%; Score 20; DB 1; Length 17;

Best Local Similarity 60.0%; Pred. No. 1.5e+03;

Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 9 GKMP 13

DB 12 GKRP 16

RESULT 8

LYC\_FELCA STANDARD; PRT: 20 AA.

ID LYC\_FELCA

AC P37153; 01-OCT-1994 (Rel. 30, Created)

DT





FT DISULFID 10 21 HYDROXYLATION.  
 FT MOD\_RES 6 7 HYDROXYLATION.  
 FT MOD\_RES 17 17 HYDROXYLATION.  
 FT MOD\_RES 22 22 AMIDATION.  
 FT HELIX 13 16  
 FT TURN 19 21  
 SQ SEQUENCE 22 AA; 2568 MW; F6CB02ADB359813C CRC64;

Query Match 27.4%; Score 20; DB 1; Length 22;  
 Best Local Similarity 60.0%; Pred. No. 1.9e+03;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 KCDER 8  
 11:1  
 DB 9 KCKDR 13

RESULT 14  
 CXM2\_CONGE STANDARD; PRT; 22 AA.  
 ID CXM2\_CONGE  
 AC P01524;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-NOV-1987 (Rel. 35, Last annotation update)  
 DE MU-CONOTOXIN GIIIB (MYOTOXIN II) (GEOGRAPHOTOXIN II) (GTX-II).  
 OS Conus geographus (Geography cone).  
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
 CC Neogastropoda; Conoidea; Conidae; Conus.  
 RN [1]  
 RN [2]  
 RN [3]  
 RP MEDLINE-85261316; PubMed-2410412;  
 RA Cruz L.J., Gray W.R., Olivera B.M., Zeikus R.D., Kerr L.,  
 RA Yoshikami D., Moczydlowski E.;  
 RT "Conus geographus toxins that discriminate between neuronal and  
 RT muscle sodium channels.";  
 RL J. Biol. Chem. 260:9280-9288(1985).  
 RN [2]  
 RN [3]  
 RP MEDLINE-83210170; PubMed-6852238;  
 RA Sato S., Nakamura H., Ohizumi Y., Kobayashi J., Hirata Y.;  
 RT "The amino acid sequences of homologous hydroxyproline-containing  
 RT myotoxins from the marine snail Conus geographus venom.";  
 RL FEBS Lett. 155:277-280(1985).  
 RN [3]  
 RP REVIEW.  
 RX MEDLINE-89024586; PubMed-3052286;  
 RA Gray W.R., Olivera B.M., Cruz L.J.;  
 RT "Peptide toxins from venomous Conus snails.";  
 RL Annu. Rev. Biochem. 57:665-700(1988).  
 RN [4]  
 RN [5]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE-96280640; PubMed-8688418;  
 RA Hill J.M., Alewood P.F., Craik D.J.;  
 RT "Three-dimensional solution structure of mu-conotoxin GIIIB, a  
 RT specific blocker of skeletal muscle sodium channels.";  
 RL Biochemistry 35:8824-8835(1996).  
 CC -1- FUNCTION: MU-CONOTOXINS ACT ON MUSCLE MEMBRANES. THEY BLOCK  
 CC MUSCLE CONTRACTION BY BINDING TO THE VOLTAGE-ACTIVATED SODIUM  
 CC CHANNELS.  
 CC PIR: A01787; MXXN2.  
 DR PIR: B23579; B23579.  
 DR PDB: 1GIB; 08-NOV-96.  
 KW Sodium channel inhibitor; Hydroxylation; Amidation; Venom;  
 KW 3D-structure.  
 FT DISULFID 3 15  
 FT DISULFID 4 20  
 FT DISULFID 10 21  
 FT MOD\_RES 6 7 HYDROXYLATION.  
 FT MOD\_RES 7 7 HYDROXYLATION.  
 FT MOD\_RES 17 17 HYDROXYLATION.  
 FT MOD\_RES 22 22 AMIDATION.

SQ SEQUENCE 22 AA; 2599 MW; F50402BA93199E01 CRC64;

Query Match 27.4%; Score 20; DB 1; Length 22;  
 Best Local Similarity 60.0%; Pred. No. 1.9e+03;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 KCDER 8  
 11:1  
 DB 9 KCKDR 13

RESULT 15  
 CXM3\_CONGE STANDARD; PRT; 22 AA.  
 ID CXM3\_CONGE  
 AC P05482;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 01-AUG-1990 (Rel. 15, Last annotation update)  
 DE MU-CONOTOXIN GIIIC.  
 OS Conus geographus (Geography cone).  
 CC Eukaryota; Metazoa; Mollusca; Gastropoda;  
 CC Neogastropoda; Conoidea; Conidae; Conus.  
 RN [1]  
 RN [2]  
 RN [3]  
 RP MEDLINE-85261316; PubMed-2410412;  
 RA Cruz L.J., Gray W.R., Olivera B.M., Zeikus R.D., Kerr L.,  
 RA Yoshikami D., Moczydlowski E.;  
 RT "Conus geographus toxins that discriminate between neuronal and  
 RT muscle sodium channels.";  
 RL J. Biol. Chem. 260:9280-9288(1985).  
 RN [2]  
 RN [3]  
 RP REVIEW.  
 RX MEDLINE-89024586; PubMed-3052286;  
 RA Gray W.R., Olivera B.M., Cruz L.J.;  
 RT "Peptide toxins from venomous Conus snails.";  
 RL Annu. Rev. Biochem. 57:665-700(1988).  
 CC -1- FUNCTION: MU-CONOTOXINS ACT ON MUSCLE MEMBRANES. THEY BLOCK  
 CC MUSCLE CONTRACTION BY BINDING TO THE VOLTAGE-ACTIVATED SODIUM  
 CC CHANNELS.  
 CC PIR: C23579; C23579.  
 DR HSSP: P01524; 1GIB.  
 KW Sodium channel inhibitor; Hydroxylation; Amidation; Venom.  
 FT DISULFID 3 15  
 FT DISULFID 4 20  
 FT DISULFID 10 21  
 FT MOD\_RES 6 7 HYDROXYLATION.  
 FT MOD\_RES 7 7 HYDROXYLATION.  
 FT MOD\_RES 17 17 HYDROXYLATION.  
 FT MOD\_RES 22 22 AMIDATION.  
 SQ SEQUENCE 22 AA; 2553 MW; F50402BA92A9813C CRC64;

Query Match 27.4%; Score 20; DB 1; Length 22;  
 Best Local Similarity 60.0%; Pred. No. 1.9e+03;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 KCDER 8  
 11:1  
 DB 9 KCKDR 13

RESULT 16  
 THAN\_PODMA STANDARD; PRT; 21 AA.  
 ID THAN\_PODMA  
 AC P55788;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE THANATIN.  
 OS Podisus maculiventris.  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;



OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomidae;  
 CC Pentatomidae; Podisus.  
 RN [1]  
 RP SEQUENCE AND CHARACTERIZATION.  
 RC MEDLINE=96165346; PubMed=8577744;  
 RA Fehlbauer P., Bulet P., Chernysh S., Briand J.-P., Rousset J.-P.,  
 RT Letellier L., Hetru C., Hoffmann J.A.;  
 RT "Structure-activity analysis of thanatin, a 21-residue inducible  
 insect defense peptide with sequence homology to frog skin  
 antimicrobial peptides."  
 RT Proc. Natl. Acad. Sci. U.S.A. 93:1221-1225(1996).  
 RN [2]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=98430981; PubMed=9760181;  
 RA Mandard N., Sodano P., Labbe H., Bonmatin J.M., Bulet P., Hetru C.,  
 RT Prak M., Vovelle F.;  
 RT "Solution structure of thanatin, a potent bactericidal and fungicidal  
 insect peptide, determined from proton two-dimensional nuclear  
 magnetic resonance data."  
 RT Eur. J. Biochem. 256:404-410(1998).  
 CC -1- FUNCTION: INSECT DEFENSE PEPTIDE WITH A BROAD SPECTRUM OF ACTIVITY  
 CC AGAINST GRAM POSITIVE AND GRAM NEGATIVE BACTERIA AND FUNGI.  
 CC NO ACTIVITY AGAINST S. AUREUS, STOPS RESPIRATION IN BACTERIA BUT  
 CC DOES NOT PERMEABILIZE THEIR INNER MEMBRANES.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- DOMAIN: FOUR REGIONS ARE IMPORTANT FOR THE ACTIVITY: THE  
 CC C-TERMINAL LOOP, THE LAST THREE C-TERMINAL RESIDUES, THE  
 CC STRETCH OF SEVEN N-TERMINAL RESIDUES AND RESIDUES 8-9 ARE  
 CC NECESSARY FOR THE ANTIFUNGAL ACTIVITY AND DISPENSABLE FOR  
 CC ANTIBACTERIAL ACTIVITY.  
 CC -1- SIMILARITY: SOME TO FROG BREVININS.  
 DR PDB: 8TEV; 02-DEC-98.  
 FT Insect immunity: Antibiotic; Fungicide; 3D-structure.  
 KM DISULFID 11 18  
 SQ SEQUENCE 21 AA; 2436 MW; C073A20CF10359E6 CRC64;

Query Match 26.78; Score 19.5; DB 1; Length 21;  
 Best Local Similarity 50.0%; Pred. No. 2.2e+03;  
 Matches 5; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

OY 2 LINKDER-GK 10  
 DB 8 ILYCKRRRTGK 17

RESULT 17  
 BAIL\_EUBSP STANDARD; PRT; 16 AA.  
 AC P32371;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE BILE ACID-INDUCIBLE OPERON PROTEIN I (FRAGMENT).  
 GN BAIL.  
 OS Eubacterium sp. (strain VPI 12708).  
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 CC Eubacterium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93259945; PubMed=8491719;  
 RA Franklund C.V., Baron S.F., Hylemon P.B.;  
 RT "Characterization of the baih gene encoding a bile acid-inducible  
 NADH:flavin oxidoreductase from Eubacterium sp. strain VPI 12708."  
 RL J. Bacteriol. 175:3002-3012(1993).  
 CC -1- PATHWAY: BILE ACID 7-ALPHA-DEHYDROXYLATION PATHWAY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U57489; AAC45418.1; -.  
 DR PIR; D36912; D36912.  
 KM Bile acid catabolism.  
 FT NON\_TER 16  
 SQ SEQUENCE 16 AA; 1721 MW; B9BAAE7B8746E4F4 CRC64;

Query Match 26.08; Score 19; DB 1; Length 16;  
 Best Local Similarity 37.5%; Pred. No. 2.1e+03;  
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 5 CDERGMKI 12  
 DB 9 CDKQDELI 16

RESULT 18  
 PKR\_VIBCH STANDARD; PRT; 16 AA.  
 ID PKR\_VIBCH  
 AC P96154;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE PHOSPHOGLYCERATE KINASE (EC 2.7.2.3) (FRAGMENT).  
 GN PKR.  
 OS Vibrio cholerae.  
 CC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OGAMA 395;  
 RX MEDLINE=97136637; PubMed=8982014;  
 RA Carroll P.A., Zhao G., Boyko S.A., Winkler M.E., Calderwood S.B.;  
 RT "Identification, sequencing, and enzymatic activity of the  
 RT erythrose-4-phosphate dehydrogenase gene of Vibrio cholerae."  
 RL J. Bacteriol. 179:293-296(1997).  
 CC -1- CATALYTIC ACTIVITY: ATP + 3-PHOSPHO-D-GLYCERATE = ADP +  
 CC 3-PHOSPHO-D-GLYCEROYL PHOSPHATE.  
 CC -1- PATHWAY: SECOND STEP IN THE SECOND PHASE OF GLYCOLYSIS.  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U72152; AAC44768.1; -.  
 DR INTERPRO; IPR001576; -.  
 DR PROSITE; PS00111; PGLYCERATE KINASE; PARTIAL.  
 KM Transferase; Kinase; Glycolysis.  
 FT NON\_TER 16  
 SQ SEQUENCE 16 AA; 1789 MW; 6BEA715980B8E227 CRC64;

Query Match 26.08; Score 19; DB 1; Length 16;  
 Best Local Similarity 36.4%; Pred. No. 2.1e+03;  
 Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 LKCDERGMKI 12  
 DB 6 MIDPLAGKRV 16

RESULT 19  
 ITHB\_HIRME STANDARD; PRT; 17 AA.

```

AC P28502:
AC 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE HIRUDIN IA (FRAGMENT).
OS Hirudo medicinalis (Medicinal leech).
CC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;
CC Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
RN [1]
RP SEQUENCE.
RX MEDLINE-90005945; PubMed-2792365;
RA Schaff M., Engels J., Tripler D.;
RA "Primary structures of new 'iso-hirudins'.";
RL FEBS Lett. 255:105-110(1989).
CC -1- FUNCTION: HIRUDIN IS A POTENT THROMBIN-SPECIFIC PROTEASE
CC INHIBITOR. IT FORMS A STABLE NON-COVALENT COMPLEX WITH ALPHA-
CC THROMBIN, THEREBY ABOLISHING ITS ABILITY TO CLEAVE FIBRINOGEN.
CC -1- SIMILARITY: BELONGS TO THE HIRUDIN FAMILY.
DR PIR: S05671; S05671.
DR HSSP: P01050; 6HIR.
KW Serine protease inhibitor; Sulfatation; Multigene family.
FT NON_TER
SQ SEQUENCE 17 AA: 1877 MW; 8904C6786C301CE7 CRC64;

Query Match
Best Local Similarity 26.0%; Score 19; DB 1; Length 17;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 CDERGK 10
DB 6 CRESGQ 11

RESULT 20
ID LPRR_BACSU STANDARD; PRT; 20 AA.
AC P23053;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE TETRACYCLINE RESISTANCE LEADER PEPTIDE.
CN TETL.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89000797; PubMed-2844262;
RA Sakaguchi R., Amano H., Shishido K.;
RA "Nucleotide sequence homology of the tetracycline-resistance
RA determinant naturally maintained in Bacillus subtilis Marburg 168
RA chromsome and the tetracycline-resistance gene of B. subtilis
RA plasmid pMS1981.";
RL Biochim. Biophys. Acta 950:441-444(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-91257555; PubMed-1646143;
RA Amano H., Sakaguchi R., Shishido K.;
RA "An insertion of Escherichia coli transposable element ISK into the
RA site immediately before tetracycline-resistance determinant of
RA Bacillus subtilis chromosomal DNA fragment in cloning in E. coli.";
RL FEBS Microbiol. Lett. 63:5-8(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: X08034; CAA30826.1; -
DR EMBL: X58999; CAA41744.1; -
DR EMBL: Z99124; CAA16115.1; -
DR PIR: S04802; LEBSTT.
DR PIR: S17282; S17282.
DR SUBTILIST; Bg11051; TELL.
KW Leader peptide; Antibiotic resistance.
SQ SEQUENCE 20 AA: 2298 MW; E870F516CA23111A CRC64;

Query Match
Best Local Similarity 26.0%; Score 19; DB 1; Length 20;
Matches 2; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 3 IKCDERGK 11
DB 1 MKCNMNRV 9

RESULT 21
ID PGK_CLOPA STANDARD; PRT; 20 AA.
AC P81346;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PUTATIVE PHOSPHOGLYCERATE KINASE (EC 2.7.2.3) (CP 11) (FRAGMENT).
CN PGK.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
RN [1]
RP SEQUENCE.
RX MEDLINE-98291870; PubMed-9629918;
RA Flengsrud R., Skjeldal L.;
RA "Two-dimensional gel electrophoresis separation and N-terminal
RA sequence analysis of proteins from Clostridium pasteurianum W5.";
RL Electrophoresis 19:802-806(1998).
CC -1- CATALYTIC ACTIVITY: ATP + 3-PHOSPHO-D-GLYCERATE = ADP +
CC 3-PHOSPHO-D-GLYCEROYL PHOSPHATE.
CC -1- PATHWAY: SECOND STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
CC 5.6, ITS MW IS: 56.2 KDA.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.
DR INTERPRO: IPR001576; -
DR PROSITE: PS00111; GLYCERATE_KINASE; PARTIAL.
KW Transferase; Kinase; Glycolysis.
FT NON_TER
SQ SEQUENCE 20 AA: 2355 MW; 749E31E595C85529 CRC64;

Query Match
Best Local Similarity 26.0%; Score 19; DB 1; Length 20;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 DERGKMT 12
DB 11 DYKGRV 17

RESULT 22
ID ACE_BOVIN STANDARD; PRT; 22 AA.
AC P12820;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-1989 (Rel. 34, Last annotation update)
DE ANGIOTENSIN-CONVERTING ENZYME; SOMATIC (EC 3.4.15.1) (ACE) (DIPEPTIDYL
DE CARBOXYPEPTIDASE I) (KININASE II) (FRAGMENT).
GN DCP1 OR ACE.

```

OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE-LUNG;  
 RX MEDLINE-87128067; PubMed-3028395;  
 RA St Clair D.K., Presper K.A., Smith P.L., Stump D.C., Heath E.C.;  
 RT "Bovine angiotensin-converting enzyme: amino-terminal sequence  
 analysis and preliminary characterization of a hybridization-selected  
 primary translation product."  
 RT Biochem. Biophys. Res. Commun. 141:968-972(1986).  
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF  
 THE TERMINAL HIS-LEU. THIS RESULTS IN AN INCREASE OF THE  
 VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.  
 CC -1- CATALYTIC ACTIVITY: RELEASE OF A C-TERMINAL DIPEPTIDE,  
 OLIGOPEPTIDE-I-XAA-XBB, WHEN XAA IS NOT PRO, AND XBB IS NEITHER  
 ASP NOR GLU. CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2 (ZINC METALLOPROTEASES).  
 CC -1- CORFACTOR: BINDS TWO ZINC IONS (BY SIMILARITY).  
 CC PIR: A26376; A26376.  
 DR MEROPS: M02.001;  
 DR INTERPRO: IPR000130;  
 DR PROSITE: PS00142; ZINC\_PROTEASE; PARTIAL.  
 KM Hydrolyase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase.  
 FT NON\_TER  
 SQ SEQUENCE 22 AA; 2271 MW; 77B0AC6A91A3893F CRC64;

Query Match 26.0%; Score 19; DB 1; Length 22;  
 Best Local Similarity 57.1%; Pred. No. 2.8e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 6 DENGKMI 12  
 DB 14 DEAGAGI 20

RESULT 23  
 ID CH10\_RHOSH STANDARD; PRT; 22 AA.  
 AC P25969;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE 10 KDA CHAPERONIN (PROTEIN CPN10) (FRAGMENT).  
 GN GROES.  
 OS Rhodospirillum rubrum (Rhodospirillum rubrum).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum group;  
 OC Rhodospirillum.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE-91329399; PubMed-1678280;  
 RA Terlesky K.C., Tabita F.R.;  
 RT "Purification and characterization of the chaperonin 10 and  
 chaperonin 60 proteins from Rhodospirillum rubrum."  
 RT Biochemistry 30:8181-8186(1991).  
 CC -1- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES  
 THE ATPASE ACTIVITY OF THE LATER.  
 CC -1- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING  
 (BY SIMILARITY).  
 CC -1- INDUCTION: BY HEAT SHOCK.  
 CC -1- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.  
 DR INTERPRO: IPR001476;  
 DR PFAM: PF00166; cpn10; 1.  
 DR PROSITE: PS00681; CHAPERONIN\_CP10; PARTIAL.  
 KM Chaperone; Heat shock.  
 FT NON\_TER  
 SQ SEQUENCE 22 AA; 2619 MW; D6027D2AF356CE8 CRC64;

Query Match 26.0%; Score 19; DB 1; Length 22;

Best Local Similarity 37.5%; Pred. No. 2.8e+03;  
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 IKCDERG 10  
 DB 14 VQSDERTK 21

RESULT 24  
 ID CYTB\_THETS STANDARD; PRT; 22 AA.  
 AC P81064;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE CYTIN B CHAIN.  
 OS Theromyzon tessellatum (Leech).  
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;  
 OC Rhynchodermatida; Glossiphoniidae; Theromyzon.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE-98055698; PubMed-9395320;  
 RA Choplin V., Biffiger T.V., Stefano G.B., Hatlar I., Salzet M.;  
 RT "Amino-acid-sequence determination and biological activity of cytin,  
 a naturally occurring specific chymotrypsin inhibitor from the leech  
 Theromyzon tessellatum."  
 RT Eur. J. Biochem. 249:733-738(1997).  
 CC -1- FUNCTION: INHIBITOR OF CHYMOTRYPSIN.  
 CC -1- SUBUNIT: HETERODIMER OF AN A CHAIN AND A B CHAIN, LINKED BY A  
 DISULFIDE BOND.  
 CC -1- SIMILARITY: SOME TO THE POTATO CHYMOTRYPSIN INHIBITOR I FAMILY  
 OF SERINE PROTEASE INHIBITOR.  
 KM Serine protease inhibitor.  
 FT DISULFID 3  
 SQ SEQUENCE 22 AA; 2504 MW; B1C217D5834F97BB CRC64;

Query Match 26.0%; Score 19; DB 1; Length 22;  
 Best Local Similarity 42.9%; Pred. No. 2.8e+03;  
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 IKCDERG 9  
 DB 1 LKCEWDG 7

RESULT 25  
 ID UC08\_MAIZE STANDARD; PRT; 15 AA.  
 AC P80614;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLOGATED COLEOPTILE (SPOT 159)  
 DE (FRAGMENT).  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE-COLEOPTILE;  
 RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
 RA Permollet J.-C., Zivy M., de Vienne D.;  
 RT "The maize 2D dimensional gel protein database: towards an integrated  
 genome analysis program."  
 RT Theor. Appl. Genet. 93:997-1005(1996).  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL, THE DETERMINED PI OF THIS UNKNOWN  
 PROTEIN IS 6.4, ITS MW IS 38.8 KDA.  
 CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE  
 FAMILY.  
 DR MAIZE-2DPAGE: P80614; COLEOPTILE.  
 DR MAIZE2DB: 123934;  
 FT NON\_TER 1 1

FT NON TER 15 15  
SQ SEQUENCE 15 AA; 1785 MW; 1978B1D6AB4DDF8D CRC64;

Query Match 24.7%; Score 18; DB 1; Length 15;  
Best Local Similarity 50.0%; Pred. No. 2.9e+03;  
Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 6 DERGKMP 13  
| | | | |  
Db 2 DWRNAMP 9

Search completed: February 5, 2001, 11:34:16  
Job time: 358 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:56:07 ; Search time 265.72 seconds  
(without alignments)  
11.027 Million cell updates/sec

Title: US-09-343-406-2  
Perfect score: 125  
Sequence: 1 GMAALPRLIAFTSEHSFSLKKGAA 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP\_ARCHAEA:\*  
2: SP\_BACTERIA:\*  
3: SP\_FUNGI:\*  
4: SP\_HUMAN:\*  
5: SP\_INVERTEBRATE:\*  
6: SP\_MAMMAL:\*  
7: SP\_MHC:\*  
8: SP\_ORGANELLE:\*  
9: SP\_PHAGE:\*  
10: SP\_PLANT:\*  
11: SP\_PROTOZOA:\*  
12: SP\_VIRUS:\*  
13: SP\_VERTEBRATE:\*  
14: SP\_UNCLASSIFIED:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	25.6	25	2	026056 helicobacte
2	30	24.0	20	6	P79837 prionaliu
3	27	21.6	25	10	040354 medicago sa
4	26.5	21.2	24	7	028NS5
5	26	20.8	9	5	028NS5
6	26	20.8	15	6	09TRP0
7	26	20.8	23	4	09R314
8	26	20.8	24	4	016333
9	25	20.0	15	11	0920U3
10	25	20.0	23	5	09NBD9
11	25	20.0	23	12	087081
12	25	20.0	24	4	016061
13	25	20.0	25	11	060839
14	25	20.0	25	11	063999
15	25	20.0	25	11	064000
16	25	20.0	25	12	09YR74
17	24	19.2	15	4	016158
18	24	19.2	16	4	09UC55
19	24	19.2	18	5	09TWS7

20	24	19.2	20	13	09PS15	09ps15 oncorhynch
21	24	19.2	22	2	09K8Q3	09k8q3 bacillus ha
22	24	19.2	22	13	09W6D7	09w6d7 gallus gall
23	24	19.2	24	10	09ZS60	09zs60 lycopersico
24	23	18.4	11	4	09U0Y1	09u0y1 homo sapien
25	23	18.4	12	6	09MZS4	09mzs4 canis famli
26	23	18.4	12	10	09ZNT6	09znt6 oryza brach
27	23	18.4	16	2	048439	048439 klebsiella
28	23	18.4	16	4	09UMJ9	09umj9 homo sapien
29	23	18.4	16	4	09NMZ2	09nmz2 homo sapien
30	23	18.4	16	12	004246	004246 human para
31	23	18.4	17	2	034216	034216 pseudomonas
32	23	18.4	18	2	09ZG42	09zg42 chlamydia t
33	23	18.4	18	5	026833	026833 trypanosoma
34	23	18.4	21	4	014595	014595 homo sapien
35	23	18.4	21	11	09OV63	09ov63 rattus sp.
36	23	18.4	22	10	09SY23	09sy23 arabidopsis
37	23	18.4	22	11	09RIC1	09ric1 mus musculu
38	23	18.4	23	2	09LA84	09la84 rhodobacter
39	23	18.4	23	11	09JIN0	09jin0 mus musculu
40	23	18.4	25	2	09R5G2	09r5g2 lactobacill
41	23	18.4	25	11	09OUX1	09oux1 rattus sp.
42	23	18.4	25	12	085452	085452 moloney mur
43	22.5	18.0	20	3	013594	013594 saccharomyc
44	22	17.6	12	10	P94011	P94011 arabidopsis
45	22	17.6	15	2	09R4T0	09r4t0 rhodobacter

## ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	25 AA.
1	026056			
AC	026056			
AD	01-JAN-1998 (TREMBLrel. 05, Created)			
DT	01-JAN-1998 (TREMBLrel. 05, Last sequence update)			
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)			
DE	HYPOTHETICAL 3.1 KDA PROTEIN.			
GN	HP1528.			
OS	Helicobacter pylori (Campylobacter pylori).			
OC	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;			
OC	Helicobacter.			
OX	NCBI_TaxID=210;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-26695 / ATCC 700392;			
RX	MEDLINE-97394467; PubMed-9252185;			
RA	Tomb J.-F., White O., Kierlavage A.R., Clayton R.A., Sutton G.G.,			
RA	Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,			
RA	Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,			
RA	Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,			
RA	Mckenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,			
RA	Berg D.E., Gocayne J.D., Venter J.M., Peterson J.D., Kelley J.M.,			
RA	Cotton M.D., Weidman J.C., Bowman C., Matthey L., Wallin E.,			
RA	Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,			
RA	Venter J.C.;			
RT	"The complete genome sequence of the gastric pathogen Helicobacter			
RT	pylori."			
RL	Nature 388:539-547(1997).			
DR	EMBL; AE000651; AAD08577.1; .			
KW	TIGR; HP1528; .			
SO	Hypothetical protein.			
	SEQUENCE 25 AA; 3075 MW; E080E93A41CE396E CRC64;			
Query Match	25.6%; Score 32; DB 2; Length 25;			
Best Local Similarity	50.0%; Pred. No. 2.1e+02;			
Matches	8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;			
QY	8 LIAFTSEHSFSLKKG 23			
DB	2 LIPIFRFLDYSFKKG 17			

```

RESULT 2
ID P79837 PRELIMINARY; PRT; 20 AA.
AC P79837;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-MAY-1997 (TREMBLrel. 03, Last annotation update)
DE IDUONATE 2-SULFATASE (FRAGMENT).
OS Ptilonotus bengalensis (leopard cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Ptilonotus.
NCBI_TaxID=37029;
RN [1]
RP SEQUENCE FROM N.A.
RA Lyons L.A., Laughlin T.F., Copeland N.G., Jenkins N.A., Womack J.E.,
RA O'Brien S.J.;
RA Nat. Genet. 15:0-0(0).
RL EMBL; U81266; AAB39359.1; -.
DR EMBL; U81266; AAB39359.1; -.
FT NON_TER 1 1
FT SEQUENCE 20 AA; 2081 MW; 7D009D35F57AE228 CRC64;
SQ

```

```

Query Match 24.0%; Score 30; DB 6; Length 20;
Best Local Similarity 62.5%; Pred. No. 3.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 8 LIAFSEH 15
DB 13 IIAFASDH 20

RESULT 3
ID Q40354 PRELIMINARY; PRT; 25 AA.
AC Q40354;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE MOS GENE (REQUIRED FOR SYNTHESIS OF L-3-O-METHYL-SCYLLO-INOSAMINE
DE (3-O-MST)), 5' END AND PROMOTER REGION (FRAGMENT).
GN MOS.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Medicago.
OX NCBI_TaxID=3879;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=89057889; PubMed=2848255;
RA Murphy P.J., Heycke N., Trenz S.P., Ratet P., De Bruijn F.J.,
RA Schell J.;
RT "Synthesis of an optine-like compound, a rhicopline, in alfalfa nodules
RT is symbolically regulated."
RL Proc. Natl. Acad. Sci. U.S.A. 85:9133-9137(1988).
DR EMBL; M23347; AAA22661.1; -.
DR INTERPRO; IPR000392; -.
DR PFAM; PF00142; fer4_N1FH; 1.
FT NON_TER 25 25
FT SEQUENCE 25 AA; 2685 MW; 1F22739287210F82 CRC64;
SQ

```

```

Query Match 21.6%; Score 27; DB 10; Length 25;
Best Local Similarity 75.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 4 ALPRTIAF 11
DB 2 AAPRTIAF 9

```

```

RESULT 4
ID Q9TNS5 PRELIMINARY; PRT; 24 AA.
AC Q9TNS5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE HUMAN LEUCOCYTE ANTIGEN BETA CHAIN DR MOLECULE HLA-DRB1
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93107319; PubMed=1469092;
RA Weyand C.M., Hickok K.C., Hander G.G., Goronzy J.J.;
RT "The HLA-DRB1 locus as a genetic component in giant cell arteritis.
RT Mapping of a disease-linked sequence motif to the antigen binding site
RT of the HLA-DR molecule."
RL J. Clin. Invest. 90:2355-2361(1992).
RW MHC.
FT SEQUENCE 24 AA; 3019 MW; 4A23BF7821BACCE0 CRC64;
SQ

```

```

Query Match 21.2%; Score 26.5; DB 7; Length 24;
Best Local Similarity 41.2%; Pred. No. 1.6e+03;
Matches 7; Conservative 3; Mismatches 2; Indels 5; Gaps 1;

```

```

QY 12 TSEHSFES----LKKG 23
DB 5 TSEYDFHNDAYLOKG 21

```

```

RESULT 5
ID Q9YV82 PRELIMINARY; PRT; 9 AA.
AC Q9YV82;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE CG18219 PROTEIN (FRAGMENT).
GN CG18219.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burks K.C., Busam D.A., Butler H., Caddieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evaragelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Moberly C., Morris J., Mostrei A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Morley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003526; AAF49437.1; -;  
 DR FLYBASE: FBgn0036647; CG18219.  
 FT NON-TER 1  
 SQ SEQUENCE 9 AA; 975 MW; F31AB4472045B9C1 CRC64;

Query Match 20.8%; Score 26; DB 5; Length 9;  
 Best Local Similarity 55.6%; Pred. No. 3.7e+05;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 16 SHEFLKKGK 24  
 |||||  
 DB 1 SHEFLKNSA 9

RESULT 6  
 Q9TRP0 PRELIMINARY; PRT; 15 AA.  
 AC O9TRP0:  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE LOW MR ZONA PELLUCIDA BINDING PROTEIN (FRAGMENT).  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 OX NCBI\_TaxID=9623;  
 RN [1]  
 RP MEDLINE=92378826; PubMed=1510840;  
 RX Parry R.V., Barker P.J., Jones R.;  
 RT "Characterization of low Mr zona pellucida binding proteins from boar  
 spermatozoa and seminal plasma."  
 RL Mol. Reprod. Dev. 33:108-115(1992).  
 SQ SEQUENCE 15 AA; 1525 MW; CC800E1028CF2058 CRC64;

Query Match 20.8%; Score 26; DB 6; Length 15;  
 Best Local Similarity 44.4%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 3 AALPLLIAF 11  
 |||::|  
 DB 7 AALPSVYVF 15

RESULT 7  
 Q9R314 PRELIMINARY; PRT; 23 AA.  
 AC Q9R314:  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE DNA ADENINE METHYLASE HOMOLOG (FRAGMENT).  
 GN M. HPTI.  
 OX Helicobacter pylori (Campylobacter pylori).

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-219, AND 213;  
 RA Raudonkluene A., Berg D.E.;  
 RT "Ic5a2 segment from H. pylori (Alaska strain 219)."  
 RT Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF008929; AAC64502.1; -;  
 DR EMBL: AF008928; AAC64500.1; -;  
 KW Methyltransferase.  
 FT NON-TER 23  
 SQ SEQUENCE 23 AA; 2635 MW; 6D6CC63737422B3 CRC64;

Query Match 20.8%; Score 26; DB 2; Length 23;  
 Best Local Similarity 45.5%; Pred. No. 1.9e+03;  
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 7 RLIAFTSESH 17  
 :|||1:  
 DB 9 KLIPFIKENIH 19

RESULT 8  
 Q16333 PRELIMINARY; PRT; 24 AA.  
 AC Q16333:  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR RECEPTOR (FRAGMENT).  
 GN FLT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX Boockock C.A., Charnock-Jones D.S., Sharkey A.M., McLaren J.,  
 RT "Expression of vascular endothelial growth factor and its receptors  
 RT fit and KDR in ovarian carcinoma."  
 RL J. Natl. Cancer Inst. 87:506-516(1995).  
 DR EMBL: S77814; AAB34002.1; -;  
 FT NON-TER 1  
 SQ SEQUENCE 24 AA; 2700 MW; 993CEB9FE226A62 CRC64;

Query Match 20.8%; Score 26; DB 4; Length 24;  
 Best Local Similarity 44.4%; Pred. No. 1.9e+03;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 10 AFTSESHF 18  
 :|||:  
 DB 14 SFSSNYFHF 22

RESULT 9  
 Q9Z0U3 PRELIMINARY; PRT; 15 AA.  
 AC Q9Z0U3:  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE SNO ONCOPROTEIN (FRAGMENT).  
 GN SNO.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Vedyo C.G., Sogayar M.C.;  
 RT "Cloning of glucocorticoid-regulated sequences using equalizing CDNA  
 RI subtraction based on selective suppression of polymerase chain  
 RL reaction in rat C6/ST1 glioma phenotypic reversion."  
 DR EMBL; AF112446; to the EMBL/GenBank/DBJ databases.  
 FT NON\_TER 15  
 SQ SEQUENCE 15 AA; 1667 MW; D2F5236DE2647B80 CRC64;

Query Match 20.0%; Score 25; DB 11; Length 15;  
 Best Local Similarity 55.6%; Pred. No. 1.7e+03;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 16 SHESLKGA 24  
 DB 6 SNFSLVQGS 14

RESULT 10  
 Q9NBD9 PRELIMINARY; PRT; 23 AA.  
 AC Q9NBD9;  
 DT 01-OCT-2000 (TRENBLrel. 15, Created)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)  
 DE HEMOGLOBIN IIB (FRAGMENT).  
 GN GB28.  
 OS Chironomus nepeanensis.  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;  
 OC Chironomidae; Chironominae; Chironomus.  
 OX NCBI\_TaxID=113501;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Grunl M.C., Schenbik S.V., Almazova K.G., Blinov A., Diez J.-L.,  
 RA Belyugin G.;  
 RT "Insect globin gene polymorphisms: Intronic minisatellites and a  
 RI retroposon interrupting exon 1 of homologous globin genes in  
 RT Chironomus (Diptera)."  
 RL Gene 251:153-163(2000).  
 DR EMBL; AF250305; AAF87714.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 FT NON\_TER 23  
 SQ SEQUENCE 23 AA; 2333 MW; 3B08F7A8FAB5EF33 CRC64;

Query Match 20.0%; Score 25; DB 5; Length 23;  
 Best Local Similarity 50.0%; Pred. No. 2.7e+03;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 RLIAFTSE 14  
 DB 6 RIVGFVSE 13

RESULT 11  
 Q87081 PRELIMINARY; PRT; 23 AA.  
 AC Q87081;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE GLYCOPROTEIN C PRECURSOR (FRAGMENT).  
 GN GC.  
 OS Pseudorabies virus.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Varicellovirus.  
 OX NCBI\_TaxID=10345;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-BECKER;  
 RX MEDLINE=96070924; PubMed=7499261;  
 RA Ryan P., Edwards C.O.;  
 RT "Systematic introduction of proline in a eukaryotic signal sequence  
 RI suggests asymmetry within the hydrophobic core."  
 RL J. Biol. Chem. 270:27876-27879(1995).  
 DR EMBL; U29124; AAC54534.1; -.  
 FT SIGNAL 1  
 FT SIGNAL 22  
 FT NON\_TER 23  
 SQ SEQUENCE 23 AA; 2229 MW; B7112B05ED406F3F CRC64;

Query Match 20.0%; Score 25; DB 12; Length 23;  
 Best Local Similarity 66.7%; Pred. No. 2.7e+03;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 MAALPRLA 10  
 DB 8 MCALPALYA 16

RESULT 12  
 Q16061 PRELIMINARY; PRT; 24 AA.  
 AC Q16061;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-JAN-1999 (TRENBLrel. 09, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE TNNT1 PROTEIN (FRAGMENT).  
 GN TNNT1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=94006957; PubMed=8403232;  
 RA Novelli G., Gennarelli M., Sangiulio F., D'Agruma L., Lo Cicero S.,  
 RA Melchionda S., Dellapiocola B.;  
 RT "Isolation and cloning by a polymerase chain reaction of a genomic DNA  
 RI fragment of the human slow skeletal troponin (TNNT1) gene."  
 RL Cell Biochemistry and Function 11:187-191(1993).  
 DR EMBL; S66057; AAD13978.1; -.  
 DR EMBL; S66170; AAD13978.1; JOINED.  
 FT NON\_TER 24  
 FT NON\_TER 24  
 SQ SEQUENCE 24 AA; 2742 MW; 4A41B1A59FD34E5E CRC64;

Query Match 20.0%; Score 25; DB 4; Length 24;  
 Best Local Similarity 55.6%; Pred. No. 2.8e+03;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 16 SHESLKGA 24  
 DB 6 SHAQKRKA 14

RESULT 13  
 Q60839 PRELIMINARY; PRT; 25 AA.  
 AC Q60839;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-AUG-1998 (TRENBLrel. 07, Last annotation update)  
 DE ALPHA-1-ACID GLYCOPROTEIN-2 (FRAGMENT).  
 GN Mus musculus (Mouse).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/C; TISSUE=LIVER;



RA Merchant M.E., Keherly M.J., Milfin R.C., Papaconstantinou J.;  
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U24278; AAA91124.1; -  
 FT NON\_TER 25 25  
 SQ SEQUENCE 25 AA; 2843 MW; 8C7FED577021D463 CRC64;

Query Match 20.0%; Score 25; DB 11; Length 25;  
 Best Local Similarity 38.5%; Pred. No. 3e+03;  
 Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 8 LIAFTSEHSFSL 20  
 Db 2 LEANPEHANFTI 14

## RESULT 14

063999 PRELIMINARY; PRT; 25 AA.  
 AC 063999;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DE VASOPRESSIN (FRAGMENT).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94286576; PubMed-8016115;  
 RA Evans D.A., der Kleij A.A., Sonnemans M.A., Burbach J.P.,  
 RA "Frameshift mutations at two hotspots in vasopressin transcripts in  
 RT post-mitotic neurons.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:6059-6063(1994).  
 DR EMBL: S71426; AAB31127.1; -  
 FT NON\_TER 1 1  
 SQ SEQUENCE 25 AA; 2478 MW; 4D78687ACF623E59 CRC64;

Query Match 20.0%; Score 25; DB 11; Length 25;  
 Best Local Similarity 75.0%; Pred. No. 3e+03;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 AALPRLIA 10  
 Db 4 AALPRASA 11

## RESULT 15

064000 PRELIMINARY; PRT; 25 AA.  
 AC 064000;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DE VASOPRESSIN (FRAGMENT).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94286576; PubMed-8016115;  
 RA Evans D.A., der Kleij A.A., Sonnemans M.A., Burbach J.P.,  
 RA "Frameshift mutations at two hotspots in vasopressin transcripts in  
 RT post-mitotic neurons.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:6059-6063(1994).  
 DR EMBL: S71427; AAB31128.1; -  
 FT NON\_TER 1 1  
 SQ SEQUENCE 25 AA; 2455 MW; B1F1E5EAD310F237 CRC64;

Query Match 20.0%; Score 25; DB 11; Length 25;  
 Best Local Similarity 75.0%; Pred. No. 3e+03;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 AALPRLIA 10  
 Db 4 AALPRASA 11

## RESULT 16

09YPU4 PRELIMINARY; PRT; 25 AA.  
 ID 09YPU4  
 AC 09YPU4;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DE L1SV-C) RNA2, COMPLETE.  
 OS Lucerne transient streak virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Sobemovirus.  
 NCBI\_TaxID=12470;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-88089532; PubMed-3335832;  
 RA Abouhaidar M.G., Pallwal Y.C.;  
 RT "Nucleotide sequence of the capsid protein gene and 3' non-coding  
 region of papaya mosaic virus RNA.";  
 RL J. Gen. Virol. 69:2369-2373(1988).  
 DR EMBL: D00341; BAA00247.1; -  
 SQ SEQUENCE 25 AA; 2868 MW; DCCE463D38B3C383 CRC64;

Query Match 20.0%; Score 25; DB 12; Length 25;  
 Best Local Similarity 46.2%; Pred. No. 3e+03;  
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 9 IAFTESEHSFSLK 21  
 Db 11 VSHTESEDTSLK 23

## RESULT 17

016158 PRELIMINARY; PRT; 15 AA.  
 ID 016158  
 AC 016158;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DE C-MYC PROTEIN (FRAGMENT).  
 GN C-MYC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94238920; PubMed-8182948;  
 RA Brenscheidt U., Eick D., Kunzmann R., Martens U., Kleinhof M.,  
 RA Mertelsmann R., Herrmann F.;  
 RT "Burkitt-like mutations in the c-myc gene locus in polymphocytic  
 RT leukemia.";  
 RL Leukemia 8:897-902(1994).  
 DR EMBL: S70386; AAB30748.1; -  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1772 MW; 883FB769FC6D5A98 CRC64;

Query Match 19.2%; Score 24; DB 4; Length 15;  
 Best Local Similarity 36.4%; Pred. No. 2.5e+03;  
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 5 LPRLIAFTSEH 15

Db 1 MPLANVSFTSRN 11

RESULT 18

ID Q9UC55 PRELIMINARY; PRT: 16 AA.  
AC Q9UC55;  
DT 01-MAY-2000 (TRENBLREL. 13, Created)  
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
DT 01-JUN-2000 (TRENBLREL. 14, Last annotation update)  
DE PROTEIN C OSAKA 10 (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE-96124381; PubMed-8560401;  
RA Miyata T., Zheng Y.Z., Sakata T., Kato H.;  
RT "Protein C Osaka 10 with aberrant propeptide processing: loss of  
RT anticoagulant activity due to an amino acid substitution in the  
RT protein C precursor."  
RL Thromb. Haemost. 74:1003-1008(1995).  
SQ SEQUENCE 16 AA; 1875 MW; EBDP581F3D325244 CRC64;

Query Match 19.2%; Score 24; DB 4; Length 16;  
Best Local Similarity 38.5%; Pred. No. 2.7e+03;

Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 AFTSEHSFSLK 22

Db 4 SFLELRHSSLER 16

RESULT 19

ID Q9TWS7 PRELIMINARY; PRT: 18 AA.  
AC Q9TWS7;  
DT 01-MAY-2000 (TRENBLREL. 13, Created)  
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
DT 01-JUN-2000 (TRENBLREL. 14, Last annotation update)  
DE 44 KDA HEMOGLOBIN B1 CHAIN (FRAGMENT).  
OS Lamellibrachia sp. (Deep-sea giant tube worm).  
OC Eukaryota; Metazoa; Vestimentifera; Basibranchia; Lamellibrachida;  
OC Lamellibrachidae; Lamellibrachia.  
OX NCBI\_TaxID=6424;  
RN [1]  
RP SEQUENCE.  
RA Suzuki T., Takagi T., Ohta S.;  
RL Zool. Sci. 10:141-146(1993).  
SQ SEQUENCE 18 AA; 2192 MW; 039D56A0B814670F CRC64;

Query Match 19.2%; Score 24; DB 5; Length 18;  
Best Local Similarity 50.0%; Pred. No. 3e+03;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 13 SEHSHF 18

Db 3 SDCHX 8

RESULT 20

ID Q9PS15 PRELIMINARY; PRT: 20 AA.  
AC Q9PS15;  
DT 01-MAY-2000 (TRENBLREL. 13, Created)  
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
DT 01-OCT-2000 (TRENBLREL. 15, Last annotation update)  
DE CREATINE KINASE (EC 2.7.3.2) (FRAGMENT).  
OS Oncorhynchus kisutch (Coho salmon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8019;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-93080727; PubMed-1449598;  
RA White K.C., Babbitt P.C., Buechter D.D., Kenyon G.L.;  
RT "The principal islet of the coho salmon (*Oncorhynchus kisutch*) contains  
RT the BB isoenzyme of creatine kinase."  
RL J. Protein Chem. 11:489-494(1992).  
DR HSSP; P00563; ZCRK.  
SQ SEQUENCE 20 AA; 2425 MW; BFBCT1D6D2A86E11 CRC64;

Query Match 19.2%; Score 24; DB 13; Length 20;  
Best Local Similarity 40.0%; Pred. No. 3.4e+03;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 11 FTSEHSFSL 20

Db 2 FGNTNNFKL 11

RESULT 21

ID Q9K8Q3 PRELIMINARY; PRT: 22 AA.  
AC Q9K8Q3;  
DT 01-OCT-2000 (TRENBLREL. 15, Created)  
DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)  
DT 01-OCT-2000 (TRENBLREL. 15, Last annotation update)  
DE BH2950. PROTEIN.  
GN BH2950.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C-125 / JCM 9153;  
RA Takami H., Nakasone K.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP001517; BAB06669.1;  
SQ SEQUENCE 22 AA; 2543 MW; 12FC8AD396A81FD8 CRC64;

Query Match 19.2%; Score 24; DB 2; Length 22;  
Best Local Similarity 66.7%; Pred. No. 3.7e+03;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 18 FSLKRG 23

Db 8 FSIERG 13

RESULT 22

ID Q9W6D7 PRELIMINARY; PRT: 22 AA.  
AC Q9W6D7;  
DT 01-NOV-1999 (TRENBLREL. 12, Created)  
DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)  
DT 01-NOV-1999 (TRENBLREL. 12, Last annotation update)  
DE HYDROXYINDOLE-O-METHYLTRANSFERASE ISOFORM A (FRAGMENT).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-99067015; PubMed-9851690;  
RA Grechez-Cassiau A., Bernard M., Ladjall K., Rodriguez J.R., Volain P.;  
RT "Structural analysis of the chicken hydroxyindole-O-methyltransferase

RT gene.";  
 RL Eur. J. Biochem. 258:44-52(1998).  
 DR EMBL: AF116455; AAD23444.1; -;  
 KW Transferase; Methyltransferase.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 22 AA; 2783 MW; 35726EABF1E450C8 CRC64;

Query Match 19.2%; Score 24; DB 13; Length 22;  
 Best Local Similarity 55.6%; Pred. No. 3.7e+03;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 12 TSEHSFSL 20  
 ||:|  
 Db 7 TSOHSFADL 15

RESULT 23  
 O9ZS60 PRELIMINARY; PRT; 24 AA.  
 AC O9ZS60;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)  
 DE PHYTOCHROME B1.  
 GN PHYB1.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; asteridae; euasteride I;  
 OC Solanales; Solanaceae; Solanum.  
 OC NCBI\_taxid=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE:99084764; PubMed:9869419;  
 RA Lazaroza G.I., Kudota T., Frances S., Peters J.L., Hughes M.J.,  
 RA Brandstader J., Szell M., Matsui M., Kendrick R.E.,  
 RA Cordomier-Pratt M., Pratt L.H.;  
 RT "Characterization of tomato PHYB1 and identification of molecular  
 RT defects in four mutant alleles."  
 RL Plant Mol. Biol. 38:1137-1146(1998).  
 DR EMBL: AJ002283; CAA05295.1; -;  
 KW Phytochrome.  
 SQ SEQUENCE 24 AA; 2573 MW; A6FA32529EDEDDB5 CRC64;

Query Match 19.2%; Score 24; DB 10; Length 24;  
 Best Local Similarity 30.8%; Pred. No. 4.1e+03;  
 Matches 4; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 13 SENSFSLKGA 25  
 :||: : :  
 Db 7 TKHSYHNSOGOA 19

RESULT 24  
 O9UOJ1 PRELIMINARY; PRT; 11 AA.  
 AC O9UOJ1;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)  
 DE PRORELAXIN H1 PRECURSOR (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Garibay-Tupas J.;  
 RT "Characterization of the human H1 relaxin 5'-flanking region."  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF104934; AAD21967.1; -;  
 FT NON\_TER 11 11

SQ SEQUENCE 11 AA; 1416 MW; 40C8681BE9C729C7 CRC64;

Query Match 18.4%; Score 23; DB 4; Length 11;  
 Best Local Similarity 57.1%; Pred. No. 2.6e+03;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 LPRLIAP 11  
 :||:|  
 Db 1 MPRLIAP 7

RESULT 25  
 O9MZS4 PRELIMINARY; PRT; 12 AA.  
 ID O9MZS4;  
 AC O9MZS4;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
 DE GROWTH HORMONE (FRAGMENT).  
 GN GH.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OC NCBI\_taxid=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lantinga-van Leeuwen I.S., Mol J.A.;  
 RT "Cloning and characterization of the 5'-flanking region of the canine  
 RT growth hormone gene."  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF166119; AAF89582.1; -;  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1229 MW; 2B587268BCB45417 CRC64;

Query Match 18.4%; Score 23; DB 6; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 2.9e+03;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 MAALPR 7  
 ||| ||  
 Db 1 MAASPR 6

Search completed: February 5, 2001, 11:27:23  
 Job time: 1876 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:48:37 ; Search time 189.33 Seconds  
(without alignments)  
4.515 Million cell updates/sec

Title: US-09-343-406-2  
Perfect score: 125  
Sequence: 1 GMAALPRLIAFTSEHSFSLKGA 25

Scoring table: BLOSUM62  
Gapop 10.0, Gapept 0.5

Searched: 268485 segs, 34193795 residues

Total number of hits satisfying chosen parameters: 141170

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

A.Geneseq\_36:\*

- 1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	125	100.0	25 16 R88667	Human glutamic acid
2	125	100.0	25 16 W18848	65 kD Glutamic acid
3	82	65.6	20 16 R72279	Glutamic acid deca
4	82	65.6	20 21 Y59552	GAD55 fragment, pe
5	67	53.6	20 16 R72278	Glutamic acid deca
6	67	53.6	20 21 Y59578	GAD55 fragment, pe
7	64	51.2	20 21 Y57065	Glutamate decarbox
8	37	29.6	15 19 W85435	Helper T-cell clas
9	37	29.6	15 21 Y73095	Hepatitis C virus
10	37	29.6	20 17 W01981	Antigenic peptide
11	37	29.6	22 21 Y73093	Hepatitis C virus
12	37	29.6	24 14 R34043	HCV NS peptide DP3

13	35	28.0	19	20	Y41898
14	34	27.2	21	19	W85304
15	34	27.2	21	21	Y73094
16	34	27.2	22	14	R41115
17	33	26.4	17	7	P60744
18	33	26.4	24	13	R29626
19	31	24.8	12	20	Y05598
20	31	24.8	15	19	W56030
21	31	24.8	17	9	P81134
22	31	24.8	18	15	R49369
23	31	24.8	22	21	T54605
24	31	24.8	23	19	W50098
25	31	24.8	23	19	W50063
26	31	24.8	23	20	Y05619
27	31	24.8	25	20	Y36603
28	29	23.2	9	17	W09332
29	29	23.2	10	13	R26156
30	29	23.2	11	17	R92601
31	29	23.2	14	9	P81970
32	29	23.2	16	13	R21716
33	29	23.2	17	17	R81593
34	29	23.2	18	12	R10998
35	29	23.2	20	11	R04962
36	29	23.2	20	15	R61281
37	29	23.2	20	19	W41195
38	29	23.2	21	20	Y02874
39	29	23.2	22	18	W22807
40	29	23.2	23	15	R45666
41	29	23.2	23	15	R53013
42	29	23.2	23	15	R53017
43	29	23.2	23	15	R53021
44	29	23.2	23	18	W27401
45	28.5	22.8	21	21	Y87379

#### ALIGNMENTS

RESULT 1

R88667 standard; peptide; 25 AA.

AC R88667;

DT 05-MAR-1996 (first entry)

DE Human glutamic acid decarboxylase residues 266-290.

KW diabetes; T-cell subpopulation; detection; antigen production;

KW diagnosis; autoimmune disease.

OS Homo sapiens.

XX DE4418091-A1.

PN 27-JUL-1995.

PD 24-MAY-1994; 94DE-4418091.

PF 04-FEB-1994; 94DE-4403522.

PR 20-JAN-1994; 94DE-4401629.

XX (ENDL/) ENDL J.

PA (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Albert W, Dormair K, Endl J, Jung G, Meini E,

PI Stahl P, Schendel D;

DR WPI; 1995-264505/35.

PT Antigen-specific activated T-lymphocytes and their detection - by

PT interaction with inventive peptide(s) of peptide-MHC complexes;

PT useful in diagnosis of e.g. diabetes and autoimmune diseases

FP Doc

XX Claim 1; Page 13; 21pp; German.  
 PS  
 CC R88667 is derived from residues 266-290 of human glutamic acid  
 CC decarboxylase and specifically reacts with a T-cell sub-population  
 CC isolated from recently diagnosed Type-I diabetics. Pharmaceutical  
 CC compsns. contg. this peptide and those shown in R88668 and R76642-62,  
 CC are useful for the diagnosis of a disease or predispositions of immune  
 CC system diseases, tumours, and autoimmune diseases, including diabetes.  
 CC The peptides are able to detect specific T-cell subpopulations that  
 CC are then used for antigen prodn., e.g. by reinjection.

XX  
 SQ Sequence 25 AA;

Query Match 100.0%; Score 125; DB 16; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-15;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMAALPRLIAFTSEHSLSLKGA 25  
 DB 1 gmaalprliatfsehslsikgaa 25

RESULT 2

W18848  
 ID W18848 standard; peptide: 25 AA.

AC W18848;

DT 05-JAN-1998 (first entry)

XX 65 KD Glutamic acid decarboxylase peptide fragment VII.

XX GAD; 65 KD: human; glutamic acid decarboxylase; autoreactive; diagnosis;  
 KW insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;  
 KW predisposition; autoimmune; tumour; rheumatoid arthritis;  
 KW multiple sclerosis.

XX Synthetic.

OS  
 PN DE19526561-A1

XX 23-JAN-1997.

XX 20-JUL-1995; 95DE-1026561.

XX 20-JUL-1995; 95DE-1026561.

XX (BOER ) BOEHRINGER MANNHEIM GMBH.

PI Donle F, Endl J, Ganz M, Jung G, Kientsech-engel R;

PI Pozzilli P, Stahl P;

WPI; 1997-088254/09.

PT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -  
 PT involving intradermal admin. of auto-reactive substances

XX Claim 11; Page 10; 12pp; German.

XX W18842-70 are peptide fragments of the 65 KD human glutamic acid  
 CC decarboxylase (GAD). The fragments are autoreactive substances used for  
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is  
 CC determined by using a claimed method for diagnosis of cell-mediated  
 CC diseases or a predisposition to cell-mediated diseases, which is effected  
 CC by administering an autoreactive substance intradermally and establishing  
 CC the diagnosis on the basis of the occurrence or lack of a positive  
 CC reaction at the site of administration. The method is used for diagnosis  
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases  
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.

XX  
 SQ Sequence 25 AA;

Query Match 100.0%; Score 125; DB 18; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-15;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMAALPRLIAFTSEHSLSLKGA 25  
 DB 1 gmaalprliatfsehslsikgaa 25

RESULT 3

R72279  
 ID R72279 standard; peptide: 20 AA.

AC R72279;

DT 13-NOV-1995 (first entry)

XX Glutamic acid decarboxylase (GAD65) fragment.

XX Glutamic acid decarboxylase; GAD65; autoimmune disorders;

XX insulin-dependent diabetes mellitus; stiff man disease.

XX Homo sapiens.

XX W09507992-A.

XX 23-MAR-1995.

XX 24-AUG-1994; 94WO-US09478.

XX 17-SEP-1993; 93US-0123859.

XX (RBGC ) UNIV CALIFORNIA.

PI Clare-Salter MJ, Erlander MG, Kaufman DL, Tobin AJ;

WPI; 1995-131360/17.

XX New polypeptide fragments of glutamic acid decarboxylase - for  
 PT diagnosis and treatment of autoimmune disease, esp. insulin  
 PT dependent diabetes, also related nucleic acid, vectors,  
 PT antibodies, hybridoma(s) etc.

XX Example 11; Page 76; 100pp; English.

XX O86481 and O86482 encode R71733 and R79105, rat and human glutamic  
 CC acid decarboxylase (GAD65) respectively, from which the GAD65  
 CC fragments described in R72261-R72298 were derived. These fragments  
 CC can be used to detect autoantibodies against GAD, e.g. to diagnose  
 CC and treat GAD-related autoimmune disorders, such as insulin  
 CC dependant diabetes mellitus or stiff man disease.

XX Sequence 20 AA;

Query Match 65.6%; Score 82; DB 16; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-08;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 AFTSEHSLSLKGA 25  
 DB 1 aftsehslsikgaa 16

RESULT 4

Y59552  
 ID Y59552 standard; peptide: 20 AA.

XX Y59552;

DT 03-APR-2000 (first entry)

```

XX GAD65 fragment, peptide #19.
DE
XX
XX GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
KW Insulin dependent diabetes mellitus; stiff man disease; diagnosis;
KM therapy.
XX
OS Homo sapiens.
XX
PN US5998366-A.
XX
PD 07-DEC-1999.
XX
PF 09-APR-1997; 97US-0827618.
XX
PR 07-JUN-1995; 95US-0485725.
PR 21-SEP-1990; 90US-0586536.
PR 18-JUN-1991; 91US-0716909.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Tobin AJ, Kaufman DL, Erlander MG;
DR WPI: 2000-095930/08.
XX
XX Ameliorating glutamic acid decarboxylase associated autoimmune
PT disorders such as insulin dependent diabetes mellitus and stiff man
PT disease -
XX
PS Claim 1; Column 42; 61pp; English.
XX
CC This sequence represents a fragment of the glutamic acid decarboxylase 65
CC (GAD65) protein. The invention relates to a method of ameliorating GAD
CC associated autoimmune disorders by administering a GAD65 peptide to the
CC patient. The method can be used for ameliorating GAD associated
CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
CC and stiff man disease. GAD65 can also be useful for screening drugs that
CC alter GAD function, for generating monoclonal antibodies and in
CC immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM
CC and the diagnosis is quite easy. It is also possible to obtain much
CC larger quantities of polypeptide via recombinant techniques than are
CC available from natural sources.
CC
XX
SQ Sequence 20 AA:

Query Match 65.6%; Score 82; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 9, 1e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 10 AFTSEHSFSLKGA 25
   |||||||
Db 1 aftsehsfslkga 16

RESULT 5
R72278
ID R72278 standard; Peptide: 20 AA.
XX
AC R72278;
XX
DT 13-NOV-1995 (first entry).
XX
DE Glutamic acid decarboxylase (GAD65) fragment.
XX
KW Glutamic acid decarboxylase; GAD65; autoimmune disorders;
KM Insulin-dependent diabetes mellitus; stiff man disease.
XX
OS Homo sapiens.
XX
PN W09507992-A.
XX
PD 23-MAR-1995.

```

```

XX
XX 24-AUG-1994; 94WO-US09478.
XX
XX 17-SEP-1993; 93US-0123859.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX
XX Clare-Salazar MJ, Erlander MG, Kaufman DL, Tobin AJ;
XX
XX WPI: 1995-131360/17.
XX
XX
XX New polypeptide fragments of glutamic acid decarboxylase - for
XX diagnosis and treatment of autoimmune disease, esp. insulin
XX dependent diabetes, also related nucleic acid, vectors,
XX antibodies, hybridoma(s) etc.
XX
XX Example 11; Page 76; 100pp; English.
XX
XX 086481 and 086482 encode R7173 and R79105, rat and human glutamic
XX acid decarboxylase (GAD65) respectively, from which the GAD65
XX fragments described in R72261-R72298 were derived. These fragments
XX can be used to detect autoantibodies against GAD, e.g. to diagnose
XX and treat GAD-related autoimmune disorders, such as insulin
XX dependent diabetes mellitus or stiff man disease.
XX
XX
XX Sequence 20 AA:

Query Match 53.6%; Score 67; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 3, 7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GMAALPRIATFSE 14
   |||||||
Db 7 gmaalpriatfse 20

RESULT 6
Y59578
ID Y59578 standard; peptide: 20 AA.
XX
AC Y59578;
XX
XX 03-APR-2000 (first entry)
XX
XX GAD65 fragment, peptide #18.
XX
XX GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
KW Insulin dependent diabetes mellitus; stiff man disease; diagnosis;
KM therapy.
XX
OS Homo sapiens.
XX
PN US5998366-A.
XX
PD 07-DEC-1999.
XX
PF 09-APR-1997; 97US-0827618.
XX
PR 07-JUN-1995; 95US-0485725.
PR 21-SEP-1990; 90US-0586536.
PR 18-JUN-1991; 91US-0716909.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX
XX Tobin AJ, Kaufman DL, Erlander MG;
XX
XX WPI: 2000-095930/08.
XX
XX Ameliorating glutamic acid decarboxylase associated autoimmune
XX disorders such as insulin dependent diabetes mellitus and stiff man
XX disease -
XX

```

PS Example 11: Column 42; 61pp; English.

XX This sequence represents a fragment of the glutamic acid decarboxylase 65  
CC (GAD65) protein. The invention relates to a method of ameliorating GAD  
CC associated autoimmune disorders by administering a GAD65 peptide to the  
CC patient. The method can be used for ameliorating GAD associated  
CC autoimmune disorders such as IDDM (Insulin dependent diabetes mellitus)  
CC and stiff man disease. GAD65 can also be useful for screening drugs that  
CC alter GAD function, for generating monoclonal antibodies and in  
CC immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM  
CC and the diagnosis is quite easy. It is also possible to obtain much  
CC larger quantities of polypeptide via recombinant techniques than are  
CC available from natural sources.

XX Sequence 20 AA;

SQ

Query Match 53.6%; Score 67; DB 21; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3,7e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GMAALPRIATFSE 14  
DB 7 gmaalpriatfse 20

RESULT 7  
Y57065  
ID Y57065 standard; peptide; 20 AA.

XX Y57065;  
AC  
XX 28-FEB-2000 (first entry)  
DT  
XX  
XX Glutamate decarboxylase peptide GAD p18.  
DE  
XX  
XX Neglected target tissue antigen; NTRA; autoimmunity; autoimmune response;  
KW Immunotherapeutic agent; Insulin dependent diabetes mellitus;  
KW Multiple sclerosis; autoimmune thyroiditis; rheumatoid arthritis;  
KW Uveoretinitis; inflammatory response.

XX Homo sapiens.  
OS  
XX  
XX W09956763-A1.  
PN  
XX 11-NOV-1999.  
PD  
XX  
XX 07-MAY-1999; 99WO-US10250.  
PF  
XX  
XX 07-MAY-1998; 98US-0084636.  
PR  
XX  
XX (REGC ) UNITV CALIFORNIA.  
PA  
XX  
XX Kaufman DL, Tian J, Olcott A;  
PI  
XX  
XX WPI: 2000-052905/04.  
DR  
XX  
XX  
PT Administration of neglected target tissue antigens to modulate immune  
PT responses -  
XX  
XX  
XX Disclosure; Page 23; 79pp; English.

XX Amino acid sequences Y57063-Y57091 are examples of neglected target  
CC tissue antigens NTRA. NTRA are antigens (whole antigens or fragments)  
CC not involved in autoimmunity. These peptides and proteins are used in the  
CC method of the invention which involves administering an NTRA as an  
CC antigen based immunotherapeutic agent, to a host afflicted with an  
CC autoimmune response associated with an autoimmune disease. The  
CC immunotherapeutic agent is used to treat autoimmune diseases such as  
CC insulin dependent diabetes mellitus, multiple sclerosis, autoimmune  
CC thyroiditis, uveoretinitis, rheumatoid arthritis or abnormal inflammatory  
CC immune responses. The NTRA induces regulatory tolerance by elicitation of  
CC regulatory T cells among T cells recognizing the NTRA but not

CC participating in the immune response. The NTRA are capable of recognition  
CC by substantial populations of uncommitted T cells which can be primed, or  
CC biased, towards regulatory responses to provide effective treatment. The  
CC NTRA are effective in regulating undesirable immune responses even when  
CC target determinants used as agents promoting tolerance agents have failed  
CC to induce an effective regulatory T cell response. NTRA as agents  
CC promoting tolerance are anticipated to be safer than use of target  
CC determinants.

XX Sequence 20 AA;

SQ

Query Match 51.2%; Score 64; DB 21; Length 20;  
Best Local Similarity 92.9%; Pred. No. 0.00012;  
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GMAALPRIATFSE 14  
DB 7 gmaavpriatfse 20

RESULT 8  
W85435  
ID W85435 standard; peptide; 15 AA.

XX W85435;  
AC  
XX  
XX 16-FEB-1999 (first entry)  
DT  
XX  
XX Helper T-cell class II peptide derived from NS4 protein.  
DE  
XX  
XX Helper T-cell peptide; human leucocyte antigen; HLA: DR4w4, DR1;  
KW DR7; cytotoxic T lymphocyte; CTL; hepatitis; autoimmune disease;  
KW Acquired immune deficiency syndrome; malaria; cancer;  
KW allograft rejection; allergy; Lyme disease; hepatitis;  
KW post-streptococcal endocarditis; glomerulonephritis;  
KW food hypersensitivity.

XX Synthetic.  
OS  
XX Hepatitis C virus.  
OS  
XX  
XX W09832456-A1.  
PN  
XX  
XX 30-JUL-1998.  
PD  
XX  
XX 23-JAN-1998; 98WO-US01373.  
PF  
XX  
XX 07-FEB-1997; 97US-0037432.  
PR  
XX 23-JAN-1997; 97US-0036713.  
PR  
XX  
XX (EPIM-) EPIMUNE INC.  
PA  
XX  
XX Sette A, Sidney J, Southwood S;  
PI  
XX  
XX WPI: 1998-427679/36.  
DR  
XX  
XX  
PT Composition containing peptide that induces cytotoxic T lymphocyte  
PT response, and helper peptide - can bind to human leucocyte antigen  
PT alleles, used to treat or prevent cancers, parasitic infections and  
PT autoimmune disease  
XX  
XX  
XX Disclosure; Page 42; 51pp; English.

XX W85284-4s1 represent helper T-cell class II peptides, which can bind to  
CC the human leucocyte antigens (HLA) DR4w4, DR1 and DR7. The peptides  
CC are used in the course of the invention. The specification describes  
CC peptides that induce a cytotoxic T lymphocyte (CTL) response, and  
CC T-helper peptides, that are used together to generate a CTL response for  
CC the treatment or prevention of viral, fungal, bacterial or parasitic  
CC infections (e.g. hepatitis, acquired immune deficiency syndrome or  
CC malaria) or cancer (e.g. renal or cervical carcinoma, lymphoma, prostate  
CC cancer or condyloma acuminatum). Helper T-cell peptides may be used  
CC alone to induce a helper T cell response, e.g. in cases of autoimmune



CC disease, allograft rejection, allergy, Lyme disease, hepatitis,  
 CC post-streptococcal endocarditis, glomerulonephritis and food  
 CC hypersensitivity.  
 XX  
 SQ Sequence 15 AA;

Query Match 29.6%; Score 37; DB 19; Length 15;  
 Best Local Similarity 61.5%; Pred. No. 4.2;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 7 RLIAFTSESHFS 19  
 ||||| : | |  
 Db 3 RLIAFTSgrnhvs 15

RESULT 9  
 ID Y73095 standard; Peptide: 15 AA.  
 XX Y73095;  
 AC  
 XX 28-FEB-2000 (first entry)  
 DT  
 XX Hepatitis C virus (HCV)-derived MHC class II (HTL) epitope, #253.  
 DE  
 XX Chimeric; pan DR epitope; expression vector;  
 KW promoter; major histocompatibility complex; MHC; targeting; peptide;  
 KW epitope; antigen; presentation; class I; cytosolic pathway;  
 KW endoplasmic reticulum; class II; extracellular antigen;  
 KW endocytic pathway; helper T lymphocyte; HTL; universal epitope;  
 KW cytotoxic T lymphocyte; CTL; immune response; immunogenicity; assay;  
 KW vaccine; immunity; infection; pathogen; virus; HIV; HBV; HCV;  
 KW hepatitis B; hepatitis C; bacterium; protozoan; tumour cell;  
 KW autoimmune disease; activation; antiviral; antimalarial;  
 KW immunoprotective.  
 KM  
 XX Synthetic.  
 OS Hepatitis C virus.  
 OS  
 XX MO9958658-A2.  
 PN  
 XX 18-NOV-1999.  
 PD  
 XX 13-MAY-1999; 99WO-US10646.  
 PF  
 XX 13-MAY-1998; 98US-0078904.  
 PR  
 XX 15-MAY-1998; 98US-0085751.  
 PR  
 XX (EPIM-) EPIMUNE INC.  
 PA  
 XX Flakes JD, Hermanson GG, Sette A, Ishloka GY, Livingston B;  
 PI Chesnut RW;  
 PI  
 XX WPI: 2000-039103/03.  
 DR  
 XX  
 XX Expression vectors encoding major histocompatibility targeting  
 PT Sequence, used as, e.g. tumor vaccines -  
 PT  
 XX Claim 11; Page 67; 130pp; English.  
 PS  
 XX Sequences Y73087-Y73102 represent hepatitis C virus (HCV)-derived MHC  
 CC class II (HTL) epitopes which are claimed for use in the present  
 CC invention. The invention relates to a novel expression vector comprising  
 CC a promoter operably linked to a fusion gene encoding a major  
 CC histocompatibility complex (MHC) targeting sequence, and two or more  
 CC heterologous peptide epitopes. The MHC targeting sequence may be a  
 CC class I targeting sequence, which directs an MHC class I epitope to  
 CC a cytosolic pathway or to the endoplasmic reticulum, or an MHC class  
 CC II targeting sequence, which directs extracellular antigens to  
 CC enter the endocytic pathway to be processed into antigen peptides  
 CC for presentation on MHC class II molecules. The heterologous  
 CC epitopes may comprise either helper T lymphocyte (HTL) epitopes,

CC or a cytotoxic T lymphocyte (CTL) epitope and a universal HTL  
 CC epitope such as a pan DR epitope (PADRE). The vectors are useful  
 CC for stimulating an immune response in vivo, as well as for use in  
 CC assaying the human immunogenicity of a human T cell peptide epitope in  
 CC vivo in a non-human mammal. They provide a nucleic acid vaccine for  
 CC enhancing immunity against infectious pathogens, such as viruses (e.g.,  
 CC HIV, hepatitis B (HBV) and hepatitis C (HCV)) bacteria, protozoa (e.g.,  
 CC Plasmodium falciparum, the cause of malaria) and also tumour cells and  
 CC autoimmune diseases. Universal MHC class II epitopes are advantageously  
 CC combined with other MHC class I and class II epitopes to increase the  
 CC number of cells that are activated in response to a given antigen and  
 CC provide a broader population coverage of MHC-reactive alleles.

CC SQ Sequence 15 AA;

Query Match 29.6%; Score 37; DB 21; Length 15;  
 Best Local Similarity 61.5%; Pred. No. 4.2;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 7 RLIAFTSESHFS 19  
 ||||| : | |  
 Db 3 RLIAFTSgrnhvs 15

RESULT 10  
 ID W01981 standard; peptide: 20 AA.  
 XX W01981;  
 AC  
 XX 26-SEP-1996 (first entry)  
 DT  
 XX Antigenic peptide comprising HCV NS4 protein residues 1916-1935.  
 DE  
 XX Antigenic peptide: C-terminal region; hepatitis C virus; HCV;  
 KW non-A non-B; non-structural protein 4; NS4; differential diagnosis;  
 KW acute; chronic; infection; binding assay.  
 KM  
 XX Hepatitis C virus.  
 OS  
 XX MO9604300-A1.  
 PN  
 XX 15-FEB-1996.  
 PD  
 XX 28-JUL-1995; 95WO-US09599.  
 PF  
 XX 29-JUL-1994; 94US-0282758.  
 PR  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PA  
 XX Fields HA, Khudyakov YE;  
 PI  
 XX WPI: 1996-129330/13.  
 DR  
 XX  
 XX Antigenic peptide(s) binding anti-hepatitis C virus antibodies -  
 PT useful for differential diagnosis of HCV in subjects  
 PT  
 XX Claim 3; Page 27; 50pp; English.  
 PS  
 XX The present peptide is an antigenic peptide (AP), derived from the  
 CC C-terminal region of the hepatitis C virus (HCV) non-structural  
 CC protein 4 (NS4). The AP was prep'd. using Fmoc chemistry, and can be  
 CC used in a claimed method for the differential diagnosis of HCV.  
 CC One, two or three of the peptides W01980-82 are separately  
 CC contacted with an antibody (Ab) contg. sample, and the binding  
 CC between the peptide(s) and the Ab determined. The Ab contg. sample  
 CC is then contacted with the peptide W01983, and the binding between  
 CC the peptide and the Ab determined. The strength of the binding  
 CC between the peptide(s) and the Ab in the two assays is compared,  
 CC where stronger binding in the 1st assay is indicative of acute HCV  
 CC infection, and equivalent binding is indicative of chronic HCV

CC Infection.  
XX Sequence 20 AA;  
SQ

Query Match  
Best Local Similarity 29.6%; Score 37; DB 17; Length 20;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
Pred. No. 5.9;

QY 7 RLIAFTSEHSFHS 19  
||||| | : | |  
Db 3 rliatfasrghvhs 15

RESULT 11  
Y73093  
ID Y73093 standard; Peptide: 22 AA.  
XX Y73093;  
AC  
XX  
XX 28-FEB-2000 (first entry)  
DT  
XX  
XX Hepatitis C virus (HCV)-derived MHC class II (HTL) epitope, #251.  
DE  
XX  
XX Chimeric; pan DR epitope; expression vector;  
KW promoter; major histocompatibility complex; MHC; targeting; peptide;  
KW epitope; antigen; presentation; class I; cytosolic pathway;  
KW endoplasmic reticulum; class II; extracellular antigen;  
KW endocytic pathway; helper T lymphocyte; HTL; unipolar epitope;  
KW cytotoxic T lymphocyte; CTL; immune response; immunogenicity; assay;  
KW vaccine; immunity; infection; pathogen; virus; HIV; HBV; HCV;  
KW hepatitis B; hepatitis C; bacterium; protozoan; tumour cell;  
KW autoimmune disease; activation; antiviral; antimalarial;  
KW immunoprotective.  
XX  
XX  
XX Synthetic.  
OS Hepatitis C virus.  
OS  
XX  
XX WO9958658-A2.  
PN  
XX  
XX 18-NOV-1999.  
PD  
XX  
XX 13-MAY-1999; 99WO-US10646.  
PF  
XX  
XX 13-MAY-1998; 98US-0078904.  
PR  
XX  
XX 15-MAY-1998; 98US-0085751.  
PR  
XX  
XX (EPIM-) EPIMUNE INC.  
PA  
XX  
XX Files JD, Hermanson GG, Sette A, Ishioka GY, Livingston B;  
PI Chesnut RW;  
PI  
XX  
XX WPI: 2000-039103/03.  
DR  
XX  
XX Expression vectors encoding major histocompatibility targeting  
PT sequence, used as, e.g. tumor vaccines -  
PT  
XX  
XX Claim 11; Page 67; 130pp; English.  
PS  
XX  
XX Sequences Y73087-Y73102 represent hepatitis C virus (HCV)-derived MHC  
CC class II (HTL) epitopes which are claimed for use in the present  
CC invention. The invention relates to a novel expression vector comprising  
CC a promoter operably linked to a fusion gene encoding a major  
CC histocompatibility complex (MHC) targeting sequence, and two or more  
CC heterologous peptide epitopes. The MHC targeting sequence may be a  
CC class I targeting sequence, which directs an MHC class I epitope to  
CC a cytosolic pathway or to the endoplasmic reticulum, or an MHC class  
CC II targeting sequence, which directs extracellular antigens to  
CC enter the endocytic pathway to be processed into antigen peptides  
CC for presentation on MHC class II molecules. The heterologous  
CC epitopes may comprise either helper T lymphocyte (HTL) epitopes,  
CC or a cytotoxic T lymphocyte (CTL) epitope and a universal HTL  
CC epitope such as a pan DR epitope (PADRE). The vectors are useful

CC for stimulating an immune response in vivo, as well as for use in  
CC assaying the human immunogenicity of a human T cell peptide epitope in  
CC vivo in a non-human mammal. They provide a nucleic acid vaccine for  
CC enhancing immunity against infectious pathogens, such as viruses (e.g.,  
CC HIV, hepatitis B (HBV) and hepatitis C (HCV)) bacteria, protozoa (e.g.,  
CC Plasmodium falciparum, the cause of malaria) and also tumour cells and  
CC autoimmune diseases. Universal MHC class II epitopes are advantageously  
CC combined with other MHC class I and class II epitopes to increase the  
CC number of cells that are activated in response to a given antigen and  
CC provide a broader population coverage of MHC-reactive alleles.  
XX  
XX  
SQ Sequence 22 AA;

Query Match  
Best Local Similarity 29.6%; Score 37; DB 21; Length 22;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
Pred. No. 6.6;

QY 7 RLIAFTSEHSFHS 19  
||||| | : | |  
Db 10 rliatfasrghvhs 22

RESULT 12  
R34043  
ID R34043 standard; peptide: 24 AA.  
XX  
XX  
XX R34043;  
AC  
XX  
XX 22-JUL-1993 (first entry)  
DT  
XX  
XX HCV NS peptide DP3.  
DE  
XX  
XX Hepatitis C virus; non structural region; antigen; diagnosis; vaccine.  
KW  
XX  
XX Synthetic.  
OS  
XX  
XX WO9306488-A.  
PN  
XX  
XX 01-APR-1993.  
PD  
XX  
XX 16-SEP-1992; 92WO-US07865.  
PF  
XX  
XX 16-SEP-1991; 91US-0762135.  
PR  
XX  
XX 12-FEB-1992; 92US-0835717.  
PR  
XX  
XX (GENE-) GENELABS TECHNOLOGIES INC.  
PA  
XX  
XX Burk KH, Dreesman GR, Pauletti D;  
PI  
XX  
XX WPI: 1993-117737/14.  
DR  
XX  
XX Detection of hepatitis C virus antigens - using HCV reactive  
PT antibody bound to solid support and competitive HCV  
PT antigen-reporter complex  
PT  
XX  
XX Claim 27; Page 122; 190pp; English.  
PS  
XX  
XX The synthetic peptide corresponds to residues 1917-1940 of the NS4  
CC domain of HCV. The peptide may be used for the sensitive and  
CC specific detection of HCV infection. Antibodies raised against the  
CC peptide can be used for passive immuno-prophylaxis and the HCV  
CC antigens can be used in vaccines to prevent HCV infection.  
CC See also R34041-65.  
XX  
XX  
SQ Sequence 24 AA;

Query Match  
Best Local Similarity 29.6%; Score 37; DB 14; Length 24;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
Pred. No. 7.4;

QY 7 RLIAFTSEHSFHS 19



ID V73094 standard; Peptide: 21 AA.  
 XX V73094;  
 AC  
 XX 28-FEB-2000 (first entry)  
 DT  
 XX  
 DE Hepatitis C virus (HCV)-derived MHC class II (HTL) epitope, #252.  
 XX  
 XX Chimeric: pan DR epitope; expression vector;  
 KM promoter: major histocompatibility complex; MHC: targeting; peptide;  
 KM epitope; antigen; presentation; class II; cytosolic pathway;  
 KM endoplasmic reticulum; class II; extracellular antigen;  
 KM endocytic pathway; helper T lymphocyte; HTL; universal epitope;  
 KM cytotoxic T lymphocyte; CTL; immune response; immunogenicity; assay;  
 KM vaccine; immunity; infection; pathogen; virus; HIV; HBV; HCV;  
 KM hepatitis B; hepatitis C; bacterium; protozoan; tumour cell;  
 KM autoimmune disease; activation; antiviral; antimalarial;  
 KM immunoprotective.  
 XX  
 OS Synthetic.  
 XX Hepatitis C virus.  
 PN W0995658-A2.  
 XX  
 PD 18-NOV-1999.  
 XX  
 PF 13-MAY-1999; 99WO-US10646.  
 XX  
 PR 13-MAY-1998; 98US-0078904.  
 PR 15-MAY-1998; 98US-0085751.  
 XX  
 PA (EPIM-) EPIMUNE INC.  
 XX  
 PI Flakes JD, Hermanson GG, Sette A, Ishioka GY, Livingston B;  
 PI Chesnut RW;  
 XX  
 DR WPI: 2000-039103/03.  
 XX  
 PT Expression vectors encoding major histocompatibility targeting  
 PT sequence, used as, e.g. tumor vaccines -  
 XX  
 PS Claim 11; Page 67; 130pp; English.  
 XX  
 CC Sequences V73087-V73102 represent hepatitis C virus (HCV)-derived MHC  
 CC class II (HTL) epitopes which are claimed for use in the present  
 CC invention. The invention relates to a novel expression vector comprising  
 CC a promoter operably linked to a fusion gene encoding a major  
 CC histocompatibility complex (MHC) targeting sequence, and two or more  
 CC heterologous peptide epitopes. The MHC targeting sequence may be a  
 CC class I targeting sequence, which directs an MHC class I epitope to  
 CC a cytosolic pathway or to the endoplasmic reticulum, or an MHC class  
 CC II targeting sequence, which directs extracellular antigens to  
 CC enter the endocytic pathway to be processed into antigen peptides  
 CC for presentation on MHC class II molecules. The heterologous  
 CC epitopes may comprise either helper T lymphocyte (HTL) epitopes,  
 CC or a cytotoxic T lymphocyte (CTL) epitope and a universal HTL  
 CC epitope such as a pan DR epitope (PADRE). The vectors are useful  
 CC for stimulating an immune response in vivo, as well as for use in  
 CC assaying the human immunogenicity of a human T cell peptide epitope in  
 CC vivo in a non-human mammal. They provide a nucleic acid vaccine for  
 CC enhancing immunity against infectious pathogens, such as viruses (e.g.,  
 CC HIV, hepatitis B (HBV) and hepatitis C (HCV)) bacteria, protozoa (e.g.,  
 CC Plasmodium falciparum, the cause of malaria) and also tumour cells and  
 CC autoimmune diseases. Universal MHC class II epitopes are advantageously  
 CC combined with other MHC class I and class II epitopes to increase the  
 CC number of cells that are activated in response to a given antigen and  
 CC provide a broader population coverage of MHC-reactive alleles.  
 XX  
 XX Sequence 21 AA:  
 SQ

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 7 RLIAFSESH 17  
 ||||| : :  
 Db 10 RLIAFASRGN 20  
 RESULT 16  
 R4115  
 ID R4115 standard; peptide: 22 AA.  
 XX  
 AC R4115;  
 XX  
 DT 22-MAR-1994 (first entry)  
 DT  
 XX  
 XX HCV peptide NS4e.  
 DE  
 XX  
 XX Human immunodeficiency virus; HIV; hepatitis C virus; HCV;  
 KM non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;  
 KM epitope; antibody; biotin; diagnosis; detection; vaccine.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "the N-terminal comprises (A)-(B)-(X)-Y; where  
 FT B= biotin;  
 FT X= biotinylation cpd. incorporated  
 FT during synthesis;  
 FT Y= bond or linking gp(s), which  
 FT minimises steric hindrance,  
 FT where Y is not a bond it is pref. 1-10  
 FT residues of (same or different) glycine,  
 FT beta-alanine, 4-aminobutyric acid,  
 FT 5-aminovaleic acid or 6-aminohexanoic acid;  
 FT parenthesis around B and X indicate opt. presence  
 FT at the specified positions but B or X must be  
 FT present in at least one of the positions shown,  
 FT B interacts with the peptide to give a cpd.  
 FT with greater diagnostic sensitivity;  
 FT A (optional)= one or more amino acids, NH2 or  
 FT gp. which modifies the N-terminus;  
 FT Z= one or more amino acids, OH, NH2, or a  
 FT linkage involving either of these 2 gps."  
 FT  
 FT Modified-site 22 /note= "the C-terminal comprises Y-(X)-Z"  
 FT  
 FT W09318054-A.  
 PN  
 XX  
 PD 16-SEP-1993.  
 XX  
 PF 08-MAR-1993; 93WO-EP00517.  
 XX  
 PR 06-MAR-1992; 92EP-0400598.  
 XX  
 PA (INNO-) INNOGENETICS NV SA.  
 XX  
 XX De LEXS R;  
 XX  
 DR WPI: 1993-303397/38.  
 XX  
 XX New biotinylated peptide(s) corresp. to immuno-dominant  
 PT epitope(s) - with increased antigenicity, useful in antibodies  
 PT detection and vaccines against hepatitis C, HIV and HTLV  
 PT  
 XX Claim 4; Page 90-98; 133pp; English.  
 PS  
 XX Peptide compns. comprise at least one and pref. a combination of  
 CC two, three, four or more biotinylated peptides chosen from the  
 CC sequences given in R41058-R41166. The peptides represent  
 CC immunologically important regions of viral proteins and are  
 CC prepd. by solid phase peptide synthesis. The compns. are  
 CC useful for the detection of antibodies to HCV, and/or HIV,

CC and/or HTLV-I or II.  
XX  
SQ Sequence 22 AA;

Query Match 27.2%; Score 34; DB 14; Length 22;  
Best Local Similarity 63.6%; Pred. No. 22;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 7 RLIAFTSEHSH 17  
||||| : :  
DB 11 RLIAFTSEHSH 21

RESULT 17  
P60744  
ID P60744 standard; Protein; 17 AA.

XX P60744;

DT 23-AUG-1991 (first entry)

DE Sequence 2 encoded by the X-region of viruses in the HTLV family, esp.  
HTLV-I and HTLV-II.

DE Adult T-cell leukaemia; X region; immunoassay; HTLV; retrovirus;  
KW epitope; anti-X region protein antibody.

XX Human T-lymphotropic viruses.

OS NO8601834-A.

PN 27-MAR-1986.

XX 19-SEP-1985; 85MO-0904763.

PF 19-SEP-1984; 84US-0652775.

PR (REGC ) UNIV OF CALIFORNIA.

PA (SLAM/) SLAMON D J.

XX Slamon DJ, Chen IS, Cline MJ;

XX WPI; 1986-094087/14.

PT New retroviral polypeptide(s), fragments antisera and antibodies  
PT - useful for detection of retroviruses in HTLV family, e.g.

PS Claim 1; Page 23; 33pp; English.

XX The inventors claim a polypeptide having a mol. wt. of about 35kd-  
CC 42kd and having at least one of the SQ in P60743-P60746. In  
CC P60743-P60746, X in the SQ corresp. to an AA such that the SQ is  
CC substantially the same as a polypeptide expressed by an HTLV, esp.  
CC HTLV-I or HTLV-II (see P60747-P60749). Also claimed is a  
CC polypeptide having a mol. wt. of not greater than about 42kd and  
CC having an AA SQ of at least 12 AAs found within at least one of  
CC the region described as PX-IV in the X SQ of HTLV-I and the region  
CC described as PX-C in the X SQ of HTLV-II are characterized by being  
CC respectively about 40kd and 37kd. Detection of antibodies to these  
CC proteins in serum may be used to determine whether exposure to an HTLV  
CC virus has occurred.

XX Sequence 17 AA;

Query Match 26.4%; Score 33; DB 7; Length 17;  
Best Local Similarity 35.3%; Pred. No. 24;  
Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 5 LPRLIAFTSEHSHSLK 21

DB :||| :||| :|||  
1 LPRLIAFTSEHSHSLK 17

RESULT 18  
R29626  
ID R29626 standard; peptide; 24 AA.

XX R29626;

DT 05-MAY-1993 (first entry)

DE GAD peptide

XX Glutamic acid decarboxylase; auto immune disease; insulin dependent;  
KW diabetes mellitus; drug screening; antibody; stiff man syndrome.

OS Synthetic.

FT Key Location/Qualifiers

FT Misc-difference 15 /label= Thr, Glu

PN EP519469-A.

PD 23-DEC-1992.

XX 17-JUN-1992; 92EP-0110308.

PR 18-JUN-1991; 91US-0716909.

PA (REGC ) UNIV CALIFORNIA.

PI Clare-Salzier MJ, Erlander MG, Kaufman DL, Tobin AJ;

DR WPI; 1992-425701/52.

PT Glutamic acid decarboxylase isolated polypeptide(s) - useful for  
PT diagnosis and treatment of auto-immune diseases, e.g. insulin  
PT dependent diabetes mellitus, drug screening and antibody prodn.

PS Claim 5; Page 20; 35pp; English.

XX The peptide corresponds to a glutamic acid decarboxylase (GAD 65)  
CC sequence and may be used for the detection of autoantibodies to GAD  
CC 65. They can also be used for screening drugs such as those that  
CC alter GAD function and for generation of antibodies which can be  
CC used to detect GAD 65. Such methods can be used in the diagnosis  
CC and therapy of autoimmune diseases such as insulin dependent  
CC diabetes mellitus and "stiff man" syndrome. The polypeptide can be  
CC used to bind or block the continuance of an autoimmune response to  
CC GAD 65. It may also be coupled to therapeutic agents and used  
CC immunotherapeutically. See also R29625-9.

XX Sequence 24 AA;

Query Match 26.4%; Score 33; DB 13; Length 24;  
Best Local Similarity 75.0%; Pred. No. 36;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GMAALPRL 8  
||||| : :  
DB 17 gmaavpkl 24

RESULT 19  
Y05598  
ID Y05598 standard; Peptide; 12 AA.

AC Y05598;  
XX 19-JUL-1999 (first entry)

XX HIV-1 group O strain KGT008 gp160 env precursor protein antigen.  
 XX HIV-1 group O: Outlier strain: V3 hypervariable region: gp160;  
 KW envelope protein; antigen; vaccine; diagnosis; AIDS.  
 XX Human immunodeficiency virus type 1.  
 OS WO9904011-A2.  
 XX 28-JAN-1999.  
 PD 20-JUL-1998; 98WO-EP04522.  
 XX 18-JUL-1997; 97EP-0870110.  
 XX (INNO-) INNOGENETICS NV.  
 PA Delaporte E, Peeters M, Saman E, Vanden Haesevelde M;  
 PI WPI; 1999-132255/11.  
 DR New isolated HIV-1 group O strains - used to produce  
 XX polynucleotides, antigens and antibodies for use in diagnosis and in  
 PT vaccines for prevention of HIV-1 infection  
 XX Claim 1; Page 89; 162pp; English.  
 PS The present sequence is an antigen of the gp160 env precursor  
 CC protein of HIV-1 group O (outlier) strain KGT008, an isolate from  
 CC Nigiria. The invention relates to new HIV-1 group O antigens (see  
 CC Y05346-625), and the use of these antigens, or nucleic acids  
 CC encoding them (see X25154-80), in the diagnosis and prophylaxis of  
 CC AIDS. They can be used as reagents for detecting HIV-1 group O  
 CC infection and for differentiating different types of HIV-1 group O  
 CC infection. Vaccines that provide protective immunity against HIV-1  
 CC infection, in particular against HIV-1 group O infection, comprise  
 CC at least one HIV-1 type O antigen, a nucleic acid encoding such an  
 CC antigen, a virus-like particle comprising such an antigen, or an  
 CC attenuated form of an HIV-1 type O strain. The invention also  
 CC relates to new HIV-1 group O strains, mostly from patients from  
 CC Cameroon and its neighbouring countries.  
 CC  
 XX Sequence 12 AA;  
 SQ

Query Match 24.8%; Score 31; DB 20; Length 12;  
 Best Local Similarity 66.7%; Pred. No. 35;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 RLIAFTSER 15  
 : | | : | | |  
 DB 4 rllcytsv 12

RESULT 20  
 W56030  
 ID W56030 standard; peptide; 15 AA.  
 XX W56030;  
 AC  
 XX 28-JUL-1998 (first entry)  
 DT  
 XX ICH-3 p20 region immunogenic peptide.  
 DE  
 XX ICH-3: mouse; programmed cell death; septic shock; folliculogenesis;  
 KW interleukin converting enzyme; ICE; cysteine protease; cytokine;  
 KM maturation; apoptosis; sepsis; burn; trauma; immunogen.  
 XX Synthetic.  
 OS Mus sp.  
 XX  
 PN WO9806263-A1.

XX 19-FEB-1998.  
 PD 08-AUG-1997; 97WO-US13898.  
 XX 09-AUG-1996; 96US-0023937.  
 PR (FISH/) FISHMAN J A.  
 XX (MIURA/) MIURA M.  
 PA (MANG/) MANG S.  
 XX (YUAN/) YUAN J.  
 PI Fishman JA, Miura M, Wang S, Yuan J;  
 DR WPI; 1998-159183/14.  
 XX Transgenic knock-out mice containing new Ich-3 disrupted gene -  
 PT useful in, e.g. screening for compounds to treat septic shock and  
 PT defects in folliculogenesis  
 PS Claim 14; Page 69; 99pp; English.  
 XX The present sequence represents an immunogenic peptide used to produce  
 CC antibodies which will bind to the p20 region of ICH-3 from the present  
 CC invention. The present invention describes: (1) a method for modulating  
 CC programmed cell death in a cell, comprising contacting the cell with  
 CC modulating amounts of ICH-3; (2) a method for promoting pro-interleukin  
 CC (pro-IL) 1-beta processing by a cell in the presence of IL-converting  
 CC enzyme (ICE), comprising contacting the cell with ICH-3 expressed under  
 CC the control of a CMV promoter; (3) a method for stimulating synthesis of  
 CC Ich-3 gene products in a cell, comprising contacting the cell with  
 CC stimulatory amounts of lipopolysaccharide (LPS); (4) a monoclonal or  
 CC polyclonal antibody that specifically binds to ICH-3; (5) a transgenic  
 CC non-human animal, comprising a disrupted Ich-3 gene, and its progeny,  
 CC and (6) a method of protecting against lung injury in common medical  
 CC conditions, comprising utilizing inhibitors of the pathway of apoptosis.  
 CC The transgenic animal is used as model for the study of septic shock or  
 CC infection. It is also used for screening for folliculogenic compounds or  
 CC for compounds for treating sepsis or septic shock or that alter  
 CC susceptibility to infection. Inhibitors of ICH-3 are used to reduce  
 CC mortality of sepsis, burns and trauma.  
 CC  
 XX Sequence 15 AA;  
 SQ

Query Match 24.8%; Score 31; DB 19; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 13 SEHSRSLKGA 24  
 : | | | | | | | | | | | | | | | |  
 DB 1 tefkhlstryga 12

RESULT 21  
 P81134  
 ID P81134 standard; protein; 17 AA.  
 XX P81134;  
 AC  
 XX 08-OCT-1990 (first entry)  
 DT  
 XX HTLV-PX-related peptide "HTLV-PX-II".  
 DE  
 XX oncogene related peptide; HTLV-PX-related.  
 KW  
 XX synthetic.  
 OS EP253325-A.  
 XX 20-JAN-1988.  
 PD 10-JUL-1987; 87EP-0109985.  
 XX

XX 11-JUL-1986; 86JP-0164331.  
 CC (YAMA/) YAMAHARA N.  
 XX WPI: 1988-015700/03.  
 DR  
 XX New oncogene related peptide -  
 PT useful in prepn of antigen and then antibodies having specific  
 PT reactivity with cancer proteins.  
 PS Claim 22; Page 33; 53pp; English.  
 CC Prepd by solid phase or soln. methods. One of 2 HTLV-Px-related  
 CC peptides among 23 claimed oncogene-related peptides.  
 CC See also P8114-33.  
 CC  
 SO Sequence 17 AA;

Query Match 24.8%; Score 31; DB 9; Length 17;  
 Best Local Similarity 35.3%; Pred. No. 54;  
 Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 5 LPRLAFTSESHPSLK 21  
 ||| : : : : ||  
 Db 1 lprlpsftgtrsktlk 17

RESULT 22  
 R49369  
 ID R49369 standard; Protein; 18 AA.  
 AC R49369;  
 DT 16-SEP-1994 (first entry)  
 DE HLA-DR beta-chain position 1-7(sic).  
 XX  
 KW Naturally-occurring; immunomodulatory protein; human; therapy; class I;  
 KW major histocompatibility complex; class II; allotype; type I diabetes;  
 KW autoimmune disease; rheumatoid arthritis; T-cell-mediated response;  
 KW multiple sclerosis; transplant rejection; vaccine; MHC.  
 OS Homo sapiens.  
 XX  
 PN WO9404171-A.  
 PD 03-MAR-1994.  
 PF 11-AUG-1993; 93WO-US07545.  
 PR 11-AUG-1992; 92US-0925460.  
 PR 15-JUN-1993; 93US-0925460.  
 PA (HARD) HARVARD COLLEGE.  
 PI Chiez RM, Hedley ML, Stern LJ, Strominger JL, Urban RG;  
 PI Vignali DA;  
 DR WPI: 1994-082825/10.  
 XX  
 PT Novel immunomodulatory peptide(s) and nucleic acids - useful for  
 PT treatment of autoimmune diseases, transplant rejection and for  
 PT vaccination  
 PS Disclosure: Page 42; 139pp; English.  
 CC The sequences given in R49291-505 and R46981-7038 represent peptide  
 CC fragments of naturally-occurring immunomodulatory proteins. These  
 CC fragments are between 10-30 residues in length and bind to a human  
 CC major histocompatibility complex (MHC) class II allotype. These  
 CC peptides may be used for therapy of autoimmune diseases, such as

CC type I diabetes, rheumatoid arthritis and multiple sclerosis, and to  
 CC reduce transplant rejection. They may also be used for vaccination  
 CC providing an exclusively T-cell-mediated response, which can be  
 CC class I or class-II based, or both, depending on the length and  
 CC character of the immunogenic peptides.  
 CC  
 SQ Sequence 18 AA;

Query Match 24.8%; Score 31; DB 15; Length 18;  
 Best Local Similarity 30.8%; Pred. No. 57;  
 Matches 4; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 6 PRLAFTSESHPS 18  
 || : : : : ||  
 Db 5 prfleystschf 17

RESULT 23  
 Y54605  
 ID Y54605 standard; Peptide; 22 AA.  
 AC Y54605;  
 DT 28-JAN-2000 (first entry)  
 DE AV37 antigen protein sequence fragment.  
 XX  
 KW AV37 antigen; monoclonal antibody; hybridoma AV37; vaccine; avian tumour;  
 KW oncogenic avian virus; Marek's disease virus; avian leucosis virus;  
 KW Rous-associated virus; reticuloendotheliosis virus; therapy.  
 OS Gallus sp.  
 PN WO955860-A1.  
 PD 04-NOV-1999.  
 PF 22-APR-1999; 99WO-GB01238.  
 PR 29-APR-1998; 98GB-0009070.  
 PA (ANIM-) INST ANIMAL HEALTH LTD.  
 PI Burgess SC, Davison TF, Ross LON;  
 DR WPI: 2000-013437/01.  
 XX  
 PT New polypeptide, useful as a vaccine and to generate monoclonal  
 PT antibodies -  
 PS Claim 6; Page 36; 63pp; English.  
 CC This sequence is a fragment of the AV37 antigen protein of the invention.  
 CC The protein is recognised by a monoclonal antibody (MAb) secreted by the  
 CC hybridoma AV37 deposited at the European Collection of Cell Cultures  
 CC (ECACC) accession number 98030304. The polypeptide can be used to isolate  
 CC a MAb, produce a hybridoma producing the MAb, and in a composition for  
 CC use as a vaccine. The vaccine can be used against oncogenic avian  
 CC viruses, including Marek's disease virus, avian leucosis virus,  
 CC Rous-associated virus and reticuloendotheliosis virus. The vector can be  
 CC used to treat avian tumours.  
 XX  
 SQ Sequence 22 AA;

Query Match 24.8%; Score 31; DB 21; Length 22;  
 Best Local Similarity 63.6%; Pred. No. 73;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 11 FTSESHPSLK 21  
 ||| || : ||  
 Db 4 ftshscdtlk 14

```

RESULT 24
ID W50098 standard; peptide: 23 AA.
XX
AC W50098;
XX
DE 26-JUN-1998 (first entry)
XX
DE Human chorionic gonadotrophin beta-chain fragment.
XX
KW Beta-chain; human; chorionic gonadotrophin; beta-hCG; treatment;
KW prevention; wasting syndrome; viral infection; cancer;
KW chronic cardiovascular disease; chemotherapy;
KW radiation therapy.
XX
OS Homo sapiens.
XX
PN W09749721-A1.
XX
PD 31-DEC-1997.
XX
PE 24-JUN-1997; 97WO-US11448.
XX
PR 09-SEP-1996; 96US-0709933.
XX
PR 24-JUN-1996; 96US-0669675.
XX
PA (UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.
XX
PI Bryant J, Gallo RC, Lunardi-Iskandar Y;
XX
DR WPI: 1998-077106/07.
XX
PT Treating or preventing wasting syndrome - by administration of human
XX
PT chorionic gonadotrophin, beta-hCG, peptides or derivatives of these
XX
PS Example: Page 78; 126pp; English.
XX
CC The present sequence is a peptide derivative of the beta-chain of
XX
CC human chorionic gonadotrophin (beta-hCG). The peptide can be used
XX
CC to treat or prevent a wasting syndrome associated with viral
XX
CC infection, e.g. human immunodeficiency syndrome virus infection,
XX
CC cancer, chronic cardiovascular disease, chemotherapy or radiation
XX
CC therapy.
XX
SQ Sequence 23 AA;

Query Match 24.8%; Score 31; DB 19; Length 23;
Best Local Similarity 29.4%; Pred. No. 77;
Matches 5; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 2 MAALPRLIAFTSEHSNF 18
: |||:: | : |
Db 3 lpalpqvcltdprf 19

RESULT 25
ID W50063 standard; peptide: 23 AA.
XX
AC W50063;
XX
DE 26-JUN-1998 (first entry)
XX
DE Human chorionic gonadotrophin beta-chain fragment.
XX
KW Beta-chain; human; chorionic gonadotrophin; beta-hCG; inhibition;
KW human immunodeficiency virus; HIV; infection; replication;
KW Kaposi's sarcoma; haematopoiesis.
XX
OS Homo sapiens.

```

```

XX
PN W09749373-A2.
XX
PD 31-DEC-1997.
XX
PE 24-JUN-1997; 97WO-US11202.
XX
PR 09-SEP-1996; 96US-0709948.
XX
PR 24-JUN-1996; 96US-0669681.
XX
PA (UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.
XX
PI Bryant J, Gallo RC, Lunardi-Iskandar Y;
XX
DR WPI: 1998-076887/07.
XX
PT Human chorionic gonadotrophin peptide derivatives - are active in
XX
PT inhibiting, e.g. HIV infection or replication, Kaposi's Sarcoma or
XX
PT have pro-haematopoietic activity
XX
PS Example: Page 110; 174pp; English.
XX
CC The present sequence is a peptide derivative of the beta-chain of
XX
CC human chorionic gonadotrophin (beta-hCG). The peptide is active in
XX
CC inhibiting, e.g. HIV infection or replication or Kaposi's sarcoma,
XX
CC or has pro-haematopoietic activity.
XX
SQ Sequence 23 AA;

Query Match 24.8%; Score 31; DB 19; Length 23;
Best Local Similarity 29.4%; Pred. No. 77;
Matches 5; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 2 MAALPRLIAFTSEHSNF 18
: |||:: | : |
Db 3 lpalpqvcltdprf 19

```

Search completed: February 5, 2001, 10:59:09  
Job time: 632 sec



Tue Feb 6 09:04:25 2001

us-09-343-406-2.rag

Page 13



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:53:09 ; Search time 141.41 Seconds  
(Without alignments)  
12.004 Million cell updates/sec

Title: US-09-343-406-2

Sequence: 125  
1 GMAALPRLIAFTSEHSFSLKKGAA 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 segs, 67900655 residues

Total number of hits satisfying chosen parameters: 5001

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR66:\*  
2: PIR3:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	25.6	25	2	H64710
2	29	23.2	20	2	S03954
3	27	21.6	14	2	PS0249
4	26	20.8	10	2	PC2171
5	26	20.8	22	2	B39138
6	26	20.8	24	2	I73619
7	25	20.0	8	2	XGH0EU
8	25	20.0	22	2	A28563
9	25	20.0	25	2	I80178
10	25	20.0	25	2	I80179
11	24.5	19.6	20	2	S07232
12	24	19.2	15	2	A61266
13	24	19.2	20	2	A53875
14	24	19.2	22	2	A40638
15	24	19.2	25	2	J00361
16	23	18.4	14	2	A60770
17	23	18.4	15	2	PS0455
18	23	18.4	15	2	FT0084
19	23	18.4	18	2	S66627
20	23	18.4	22	2	A17267
21	23	18.4	23	2	PH0858
22	23	18.4	24	2	A36912
23	23	18.4	24	4	S09363
24	22	17.6	10	2	S74176
25	22	17.6	12	2	S01172
26	22	17.6	12	2	PH1635
27	22	17.6	13	2	B61620
28	22	17.6	15	2	PA0087
29	22	17.6	18	2	S48862

30	22	17.6	19	2	S43652	serpin MSZC1 - whe
31	22	17.6	20	2	B41299	T-cell receptor al
32	22	17.6	22	2	S59071	secretory protein
33	22	17.6	22	2	I37144	aspartylglycosamin
34	22	17.6	23	2	S51188	aldehyde dehydrog
35	22	17.6	25	2	T47238	glutamate dehydrog
36	22	17.6	25	2	A10759	isocitrate dehydro
37	22	17.6	25	2	S36378	Ig heavy chain V r
38	21.5	17.2	21	2	I49414	gene CTR1-1 protei
39	21.5	17.2	23	2	PH1641	Ig H chain V-D-J r
40	21	16.8	6	4	A35039	hypothetical colla
41	21	16.8	13	2	A38929	glutathione peroxi
42	21	16.8	13	2	H56046	urinary tract ston
43	21	16.8	16	1	LEFCH	his operon leader
44	21	16.8	16	2	C39509	mannose-specific 1
45	21	16.8	16	2	C61414	chymotrypsin (EC 3

#### ALIGNMENTS

RESULT 1  
H64710  
hypothetical protein HP1528 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: H64710  
R:TomD, J.F.; White, O.; Kervavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKee-  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey,  
Nature 386, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser,  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467  
A:Accession: H64710  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-25 <TOM>  
A:Cross-references: GB:AE006651, GB:AE000511, NID:92314708, PIDN:AD08577.1, PID:9231

Query Match 25.6%; Score 32; DB 2; Length 25;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 8 LIAFTSEHSFSLKKG 23  
|||  
DB 2 LIPFYRFIDSLKKG 17

RESULT 2  
S03954  
acidic fibroblast growth factor - pig (fragment)  
N:Alternate names: alpha-endothelial cell growth factor  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 18-Jun-1993  
C:Accession: S03954  
R:Quinkler, W.; Maasberg, M.; Bernotat-Danielowski, S.; Luethke, N.; Sharma, H.S.; Sch  
Eur. J. Biochem. 181, 67-73, 1989  
A:Title: Isolation of heparin-binding growth factors from bovine, porcine and canine  
A:Reference number: S03953; MUID:89231704  
A:Accession: S03954  
A:Molecule type: Protein  
A:Residues: 1-20 <QUI>  
C:Keywords: growth factor

Query Match 23.2%; Score 29; DB 2; Length 20;  
Best Local Similarity 35.7%; Pred. No. 2.7e+02;  
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 LPIELIAFTSEHSF 18  
|||

DB 3 LPPKLYSSNGGHF 16

RESULT 3  
PS0249

porin - rice (strain Nihonbare) (fragment)

C:Species: Oryza sativa (rice)

C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 11-Apr-1995

C:Accession: PS0249

R:Tsugita, A.

submitted to JIPID, April 1993

A:Reference number: PS0206

A:Accession: PS0249

A:Molecule type: protein

A:Residues: 1-14 <RSU>

A:Experimental source: callus

Query Match 21.6%; Score 27; DB 2; Length 14;  
Best Local Similarity 30.8%; Pred. No. 3.9e+02;

Matches 4; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 9 IATSESHSHSLK 21

DB 2 VITPDHTANGIK 14

RESULT 4

PC2171

triacylglycerol lipase (EC 3.1.1.3) I A chain - Rhizopus niveus (strain IFO 4759) (fragm

C:Species: Rhizopus niveus

C:Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 07-May-1999

C:Accession: PC2171

R:Kohno, M.; Kugimiya, W.; Hashimoto, Y.; Morita, Y.

Biochem. Biotechnol. Biochem. 58, 1007-1012, 1994

A:Title: Purification, characterization, and crystallization of two types of lipase from

A:Reference number: PC2171; MUID:94319059

A:Accession: PC2171

A:Molecule type: protein

A:Residues: 1-10 <KOH>

C:Comment: This enzyme catalyzes the hydrolysis of the ester bonds.

C:Keywords: carboxylic ester hydrolase

Query Match 20.8%; Score 26; DB 2; Length 10;

Best Local Similarity 75.0%; Pred. No. 3.9e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AALPRLIA 10

DB 3 AALPPLIS 10

RESULT 5

B39138

dihydrolipoamide S-acetyltransferase (EC 2.3.1.12) - Pelobacter carbinolicus (fragment)

C:Species: Pelobacter carbinolicus

C:Date: 27-Nov-1991 #sequence\_revision 27-Nov-1991 #text\_change 05-May-2000

C:Accession: B39138

R:Opfermann, F.B.; Schmidt, B.; Steinbuechel, A.

J. Bacteriol. 173, 757-767, 1991

A:Title: Purification and characterization of acetoin:2,6-dichlorophenolindophenol oxid

A:Reference number: A39138; MUID:91100366

A:Accession: B39138

A:Molecule type: preliminary

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-22 <OPP>

C:Superfamily: dihydrolipoamide acetyltransferase; lipoyl-biotin-binding homology

C:Keywords: acyltransferase; coenzyme A

Query Match 20.8%; Score 26; DB 2; Length 22;

Best Local Similarity 29.4%; Pred. No. 9.1e+02;

Matches 5; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 7 RLIAFTSESHSHSLKKG 23

DB 4 RLIALTPKKGLTWESG 20

RESULT 6

I73619

endothelial growth factor receptors flt [imported] - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 21-Jul-2000

C:Accession: I73619

R:Boockock, C.A.; Charnock-Jones, D.S.; Sharkey, A.M.; McLaren, J.; Barker, P.J.; Wrig

J. Natl. Cancer Inst. 87, 506-516, 1995

A:Title: Expression of vascular endothelial growth factor and its receptors flt and K

A:Reference number: I56493; MUID:95222657

A:Accession: I73619

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-24 <RES>

A:Cross-references: GB:S77814; NID:998566; PIDN:AAB34002.1; PTD:998567

C:Genetics:

A:Gene: flt

Query Match 20.8%; Score 26; DB 2; Length 24;

Best Local Similarity 44.4%; Pred. No. 9.9e+02;

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 10 APTSESHSHF 18

DB 14 SFSNMYFHF 22

RESULT 7

XGHDEU

urine glycopeptide - human

C:Species: Homo sapiens (man)

C:Date: 20-Jun-2000 #sequence\_revision 20-Jun-2000 #text\_change 20-Jun-2000

C:Accession: A03188

R:Lotz, C.J.; Weiss, J.B.

Biochem. J. 123, 25P, 1971

A:Title: Identification in urine of a low-molecular-weight polar glycopeptide contain

A:Reference number: A03188; MUID:72062338

A:Accession: A03188

A:Molecule type: protein

A:Residues: 1-8 <LOT>

C:Comment: The identity of the glycoprotein from which this peptide is derived is unk

re has also been found (see PIR:XGHDE).

C:Superfamily: unassigned animal peptides

C:Keywords: glycoprotein

F:1/Binding site: carbohydrate (Cys) (covalent) #status experimental

Query Match 20.0%; Score 25; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 2e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 EHSF 17

DB 2 EHSF 5

RESULT 8

A28563

hemoglobin chain III - earthworm (Lumbricus terrestris) (fragment)

C:Species: Lumbricus terrestris (common earthworm)

C:Date: 19-Nov-1988 #sequence\_revision 19-Nov-1988 #text\_change 04-Mar-2000

C:Accession: A28563

R:Gotch, T.; Shishikura, F.; Snow, J.W.; Erefej, K.I.; Vlnogradov, S.N.; Walz, D.A.

Biochem. J. 241, 441-445, 1987

A:Title: Two globin strains in the giant annelid extracellular haemoglobins.  
 A:Reference number: A90337; MUID:87241210  
 A:Accession: A28563  
 A:Molecule type: protein  
 A:Residues: 1-22 <GOT>  
 C:Superfamily: globin, globin homology  
 C:Keywords: oxygen carrier

Query Match 20.0%; Score 25; DB 2; Length 22;  
 Best Local Similarity 40.0%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 13 SEHSFSLK 22  
 ||| : : :  
 Db 8 SEEDHYIVK 17

RESULT 9  
 180178  
 vasopressin - rat (fragment)  
 C:Species: Rattus sp. (rat)  
 C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
 C:Accession: I80178  
 R:Evans, D.A.; van der Kleij, A.A.; Sonnemans, M.A.; Burdach, J.P.; van Leeuwen, F.W.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 6059-6063, 1994  
 A:Title: Frameshift mutations in vasopressin transcripts in post-mitotic  
 A:Reference number: I59310; MUID:94286576  
 A:Accession: I80178  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-25 <RGS>  
 A:Cross-references: GB:S71426; NID:g557990; PIDN:AAB31127.1; PID:g557991

Query Match 20.0%; Score 25; DB 2; Length 25;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+03;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AALPRLIA 10  
 ||||| :  
 Db 4 AALPRASA 11

RESULT 10  
 180179  
 vasopressin - rat (fragment)  
 C:Species: Rattus sp. (rat)  
 C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
 C:Accession: I80179  
 R:Evans, D.A.; van der Kleij, A.A.; Sonnemans, M.A.; Burdach, J.P.; van Leeuwen, F.W.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 6059-6063, 1994  
 A:Title: Frameshift mutations at two hotspots in vasopressin transcripts in post-mitotic  
 A:Reference number: I59310; MUID:94286576  
 A:Accession: I80179  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-25 <RES>  
 A:Cross-references: GB:S71427; NID:g557992; PIDN:AAB31128.1; PID:g557993

Query Match 20.0%; Score 25; DB 2; Length 25;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+03;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AALPRLIA 10  
 ||||| :  
 Db 4 AALPRASA 11

RESULT 11  
 S07232  
 ribulose-bisphosphate carboxylase subunit-binding protein alpha chain - garden pea (frag

N:Alternate names: rubisco subunit-binding protein alpha chain  
 C:Species: Pisum sativum (garden pea)  
 C>Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 12-Sep-1997  
 C:Accession: S07232  
 R:Mugrove, J.E.; Johnson, R.A.; Ellis, R.J.  
 Eur. J. Biochem. 163, 529-534, 1987  
 A:Title: Dissociation of the ribulosebisphosphate-carboxylase large-subunit binding P  
 A:Reference number: S07232; MUID:87161853  
 A:Accession: S07232  
 A:Molecule type: protein  
 A:Residues: 1-20 <MUS>  
 C:Comment: This protein binds the newly synthesized large subunit and the newly impor  
 C:Superfamily: chaperonin groEL  
 C:Keywords: chloroplast; heterododecamer; molecular chaperone

Query Match 19.6%; Score 24.5; DB 2; Length 20;  
 Best Local Similarity 40.0%; Pred. No. 1.4e+03;  
 Matches 6; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 9 IAFTEHSFSLK 23  
 ||| : : :  
 Db 5 IAF-DQHSRSMQAG 18

RESULT 12  
 A61266  
 glucuronosyltransferase (EC 2.4.1.17), 4-hydroxybiphenyl-specific - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 12-May-1994 #sequence\_revision 12-May-1994 #text\_change 13-Sep-1998  
 C:Accession: A61266  
 R:Styczynski, P.; Green, M.; Pulg, J.; Coffman, B.; Tepfly, T.  
 Mol. Pharmacol. 40, 80-84, 1991  
 A:Title: Purification and properties of a rat liver phenobarbital-inducible 4-hydroxy  
 A:Reference number: A61266; MUID:91312327  
 A:Accession: A61266  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-15 <STR>  
 C:Superfamily: glucuronosyltransferase  
 C:Keywords: glycoprotein; glucosyltransferase; hexosyltransferase

Query Match 19.2%; Score 24; DB 2; Length 15;  
 Best Local Similarity 25.0%; Pred. No. 1.3e+03;  
 Matches 3; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 7 RLIFTEHSHF 18  
 : : : : : :  
 Db 2 KVLVPMPEYSHW 13

RESULT 13  
 A53875  
 creatine kinase (EC 2.7.3.2) CK-MM - coho salmon (fragment)  
 C:Species: Oncorhynchus kisutch (coho salmon)  
 C>Date: 26-Sep-1994 #sequence\_revision 18-Nov-1994 #text\_change 11-Apr-1997  
 C:Accession: A53875  
 R:White, R.C.; Babbitt, P.C.; Buechter, D.D.; Kenyon, G.L.  
 J. Protein Chem. 11, 489-494, 1992  
 A:Title: The principal islet of the coho salmon (Oncorhynchus kisutch) contains the BB  
 A:Reference number: A53875; MUID:93080727  
 A:Accession: A53875  
 A>Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-20 <WHT>  
 A:Experimental source: Brockmann body, principal islet  
 A:Note: sequence extracted from NCBI backbone (NCBI:P120599)  
 C:Superfamily: creatine kinase; creatine kinase repeat homology  
 C:Keywords: phosphotransferase

Query Match 19.2%; Score 24; DB 2; Length 20;  
Best Local Similarity 40.0%; Pred. No. 1.7e+03;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 11 FTSEHSESL 20  
| : | : | : |  
Db 2 FGNTHNNFKL 11

## RESULT 14

540638  
ATF-43 protein - human (fragments)

C:Species: Homo sapiens (man)  
C:Date: 19-May-1994 #sequence\_revision 10-Nov-1995 #text\_change 10-Nov-1995  
C:Accession: S40638  
R:Hurst, H.C.; Totty, N.F.; Jones, N.C.  
Nucleic Acids Res. 19, 4601-4609, 1991  
A:Title: Identification and functional characterisation of the cellular activating trans  
A:Reference number: S40638; M0ID:91367654  
A:Accession: S40638  
A>Status: Preliminary  
A:Molecule type: protein  
A:Residues: 1-22 <HUR>

Query Match 19.2%; Score 24; DB 2; Length 22;  
Best Local Similarity 40.0%; Pred. No. 1.9e+03;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 ALPRLIAFTS 13  
| : | : | : |  
Db 13 SUPQIVVMTS 22

## RESULT 15

300361  
vasoactive intestinal peptide - Atlantic cod (fragment)

C:Species: gadus morhua (Atlantic cod)  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 21-Nov-1997  
C:Accession: J00361  
R:Thwaites, D.T.; Young, J.; Thorndyke, M.C.; Dimaline, R.  
Regul. Pept. 21, 436, 1988  
A:Title: Isolation and characterisation of two teleost VIP's.  
A:Reference number: J00361  
A:Accession: J00361  
A:Molecule type: protein  
A:Residues: 1-25 <THM>  
C:Superfamily: glucagon  
C:Keywords: duplication; intestine; neuropeptide

Query Match 19.2%; Score 24; DB 2; Length 25;  
Best Local Similarity 35.7%; Pred. No. 2.2e+03;  
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 11 FTSEHSESLKGA 24  
| : | : | : |  
Db 6 FTDNYSFRKQMAA 19

## RESULT 16

A60770  
cell surface alloantigen gp60 - human (fragment)

C:Species: Homo sapiens (man)  
C:Date: 14-May-1993 #sequence\_revision 14-May-1993 #text\_change 07-Feb-1997  
C:Accession: A60770  
R:Schandorf, D.; Yamaguchi, H.; Old, L.J.; Srivastava, P.K.  
J. Immunol. 142, 1621-1625, 1989  
A:Title: A novel heteromorphic human cell surface alloantigen, gp60, defined by a human  
A:Reference number: A60770; M0ID:89140352  
A:Accession: A60770  
A:Molecule type: protein  
A:Residues: 1-14 <SCH>

C:Comment: This protein is an alloantigen in human populations but is not found on ce  
C:Comment: This protein exists in both membrane bound and cytosolic forms.  
C:Keywords: glycoprotein; polymorphism; surface antigen

Query Match 18.4%; Score 23; DB 2; Length 14;  
Best Local Similarity 40.0%; Pred. No. 1.7e+03;  
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 IAFTEHSHF 18  
| : | : | : |  
Db 5 VSFTKESXGF 14

## RESULT 17

PS0455  
superoxide dismutase (EC 1.15.1.1) (Mn) - rice (strain Nihonbare) (fragment)

C:Species: Oryza sativa (rice)  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 05-Mar-1999  
C:Accession: PS0455  
R:Tsuigita, A.  
submitted to JIPID, April 1993  
A:Reference number: PS0206  
A:Accession: PS0455  
A:Molecule type: protein  
A:Residues: 1-15 <TSU>  
A:Experimental source: germ  
C:Function:  
A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxy  
C:Keywords: metalloprotein; oxidoreductase

Query Match 18.4%; Score 23; DB 2; Length 15;  
Best Local Similarity 62.5%; Pred. No. 1.8e+03;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GMAALPRL 8  
| : | : | : |  
Db 2 GVALAPRL 9

## RESULT 18

PT0094  
succinate dehydrogenase (ubiquinone) (EC 1.3.5.1) 27k iron-sulfur protein, mitochondr

C:Species: Mus musculus (house mouse)  
C:Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 21-Aug-1998  
C:Accession: PT0094  
R:Kawakami, T.; Uchida, T.; Sakai, T.; Kano, M.; Morimasa, T.; Tsugita, A.  
submitted to JIPID, July 1998  
A:Description: Proteome analysis of mouse brain.  
A:Reference number: PT0091  
A:Accession: PT0094  
A:Molecule type: protein  
A:Residues: 1-15 <KAW>  
A:Experimental source: brain, striatum  
C:Keywords: mitochondrion; oxidoreductase

Query Match 18.4%; Score 23; DB 2; Length 15;  
Best Local Similarity 55.6%; Pred. No. 1.8e+03;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 AALPRLIAF 11  
| : | : | : |  
Db 6 AAAPRIKXF 14

## RESULT 19

S66627  
tau protein - human (fragments)

C:Species: Homo sapiens (man)  
C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 13-Mar-1997  
C:Accession: S66627

R:Moreno, F.J.; Medina, M.; Perez, M.; Montejo de Garcini, E.; Avila, J.  
 FEBS Lett. 372, 65-68, 1995  
 A:Title: Glycogen synthase kinase 3 phosphorylates recombinant human tau protein at ser1  
 A:Reference number: S66627; MUID:96032547  
 A:Accession: S66627  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-8;9-18 <MOR>

Query Match 18.4%; Score 23; DB 2; Length 18;  
 Best Local Similarity 62.5%; Pred. No. 2.2e+03;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 SHFSLKRG 23  
 | | | | |  
 Db 3 STFNLRKG 10

RESULT 20

complement C4 - guinea pig (fragment)  
 C:Species: Cavia porcellus (guinea pig)  
 C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 30-Sep-1993  
 C:Accession: A17267  
 R:Goldberger, G.; Abraham, G.N.; Williams, J.; Colten, H.R.  
 J. Biol. Chem. 255, 7071-7074, 1980  
 A:Title: NH2-terminal sequence analysis of pro-C4, the precursor of the fourth component  
 A:Reference number: A17267; MUID:80227885  
 A:Accession: A17267  
 A:Molecule type: protein  
 A:Residues: 1-22 <GOL>  
 A:Note: 18-Val was also found  
 C:Superfamily: alpha-2-macroglobulin

Query Match 18.4%; Score 23; DB 2; Length 22;  
 Best Local Similarity 66.7%; Pred. No. 2.8e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 PRLTAF 11  
 | | | | |  
 Db 2 PRLTAF 7

RESULT 21

MAUD protein - Paracoccus denitrificans (fragment)  
 C:Species: Paracoccus denitrificans  
 C:Date: 17-Aug-1992 #sequence\_revision 17-Aug-1992 #text\_change 08-Oct-1999  
 C:Accession: PH0858  
 R:Christoserdov, A.Y.; Boyd, J.; Mathews, F.S.; Lidstrom, M.E.  
 Biochem. Biophys. Res. Commun. 184, 1181-1189, 1992  
 A:Title: The genetic organization of the mau gene cluster of the facultative autotroph F  
 A:Reference number: PH0856; MUID:92272706  
 A:Accession: PH0858  
 A:Molecule type: DNA  
 A:Residues: 1-23 <CHT>  
 A:Cross-references: GB:M90098; NID:9150580; PIDN:AAA25577.1; PID:9150581  
 C:Genetics:

Query Match 18.4%; Score 23; DB 2; Length 23;  
 Best Local Similarity 36.0%; Pred. No. 2.9e+03;  
 Matches 9; Conservative 3; Mismatches 5; Indels 8; Gaps 1;

QY 1 GMAALPRLIAFTSEHSHSLKGA 25  
 | | | | |  
 Db 7 GFASLQGYMA-----SRKQAA 23

RESULT 22

A36912  
 hypothetical protein 1 batG-region [imported] - Eubacterium sp. (fragment)  
 C:Species: Eubacterium sp.  
 C:Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Jun-2000  
 C:Accession: A36912  
 R:Franklund, C.V.; Baron, S.F.; Hylemon, P.B.  
 J. Bacteriol. 175, 3002-3012, 1993  
 A:Title: Characterization of the batH gene encoding a bile acid-inducible NADH:flavin  
 A:Reference number: A36912; MUID:92259945  
 A:Accession: A36912  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-24 <PRA>  
 A:Experimental source: VPI 12708  
 A:Note: sequence extracted from NCBI backbone (NCBI:131883, NCBI:131906)

Query Match 18.4%; Score 23; DB 2; Length 24;  
 Best Local Similarity 50.0%; Pred. No. 3e+03;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 LPRLIATSE 14  
 | | | | |  
 Db 4 LPVLILFTQ 13

RESULT 23

S09363  
 hypothetical MTCOL/MTCYB mutant fusion protein - human mitochondrion (fragment)  
 C:Species: mitochondrion Homo sapiens (man)  
 C:Date: 21-Nov-1993 #sequence\_revision 14-Aug-1997 #text\_change 20-Apr-2000  
 C:Accession: S09363  
 R:Poulton, J.; Deadman, M.E.; Gardiner, R.M.  
 Nucleic Acids Res. 17, 10223-10229, 1989  
 A:Title: Tandem direct duplications of mitochondrial DNA in mitochondrial myopathy: a  
 A:Reference number: S09363; MUID:90098864  
 A:Accession: S09363  
 A:Molecule type: DNA  
 A:Residues: 1-24 <POU>  
 C:Comment: This is the hypothetical translation of a sequence believed to result from  
 C:Genetics:  
 A:Gene: MTCOL/MTCYB  
 A:Genome: mitochondrion  
 A:Genetic code: SGC1  
 A:Keywords: fusion protein; mitochondrion  
 F:1-4/Region: cytochrome-c oxidase chain I  
 F:5-24/Region: cytochrome b (+2 frame shifted)

Query Match 18.4%; Score 23; DB 4; Length 24;  
 Best Local Similarity 66.7%; Pred. No. 3e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 HFSLRK 22  
 | | | | |  
 Db 7 HFSLRK 12

RESULT 24

S74176  
 glucokinase (EC 2.7.1.12), thermoresistant - Escherichia coli (fragment)  
 C:Species: Escherichia coli  
 C:Date: 14-Apr-1998 #sequence\_revision 24-Apr-1998 #text\_change 07-May-1999  
 C:Accession: S74176  
 R:Izu, H.; Adachi, O.; Yamada, M.  
 FEBS Lett. 394, 14-16, 1996  
 A:Title: Purification and characterization of the Escherichia coli thermoresistant gl  
 A:Reference number: S74176; MUID:97074194  
 A:Accession: S74176  
 A:Molecule type: protein  
 A:Residues: 1-10 <IZU>  
 A:Experimental source: strain K-12  
 C:Genetics:

A:Gene: gntK  
C:Keywords: dimer; phosphotransferase

Query Match 17.6%; Score 22; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.7e+03;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 12 TSEHSH 17  
1: 1 1  
DB 2 TTNHSH 7

## RESULT 25

S01122  
Photosystem II 3.7k protein - spinach (fragment)  
C:Species: Spinacia oleracea (spinach)  
C:Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 18-Jun-1993  
C:Accession: S01122  
R:Schroeder, W.P.; Henrysson, T.; Akerlund, H.E.  
FEBS Lett. 235, 289-292, 1988  
A:Title: Characterization of low molecular mass proteins of photosystem II by N-terminal  
A:Reference number: S01120  
A:Accession: S01122  
A:Molecule type: Protein  
A:Residues: 1-12 <SCH>  
C:Keywords: chloroplast; photosynthesis; photosystem II; thylakoid

Query Match 17.6%; Score 22; DB 2; Length 12;  
Best Local Similarity 55.6%; Pred. No. 2.1e+03;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 5 LPRLAFTS 13  
1: 1 1 1  
DB 2 LPEAYAFIS 10

Search completed: February 5, 2001, 11:10:43  
Job time: 1054 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 11:28:18 ; Search time 78.94 Seconds  
(without alignments)  
10.227 Million cell updates/sec

Title: US-09-343-406-2

Sequence: 1 GMAALPRLIAFTSEHSFLKKGAA 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 1436

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28.5	22.8	25	1	MISG_MISAN
2	26	20.8	25	1	CR11_LITSP
3	26	20.8	25	1	CR12_LITCE
4	26	20.8	25	1	CR13_LITCE
5	26	20.8	25	1	CR14_LITGI
6	25	20.0	8	1	GLUR_HUMAN
7	25	20.0	24	1	FRED_LITIN
8	24	19.2	25	1	CR15_LITCE
9	24	19.2	25	1	VIP_GADMO
10	23	18.4	22	1	CO4_CAVPO
11	22.3	18.0	23	1	CH60_THIFE
12	22	17.6	13	1	LMT4_LOEMI
13	22	17.6	18	1	CYCH_MOUSE
14	21	16.8	8	1	FAR7_ASCSU
15	21	16.8	16	1	LPHI_ECOLI
16	21	16.8	22	1	SETB_SALTY
17	21	16.8	24	1	GAB6_RANRU
18	20	16.0	8	1	ALU1_CYPPO
19	20	16.0	10	1	GLEW_HUMAN
20	20	16.0	10	1	GON1_PETMA
21	20	16.0	15	1	MCA2_RHOOP
22	20	16.0	20	1	DETS_RAT
23	20	16.0	20	1	LPRB_BACST
24	20	16.0	22	1	CYSP_TRIVA
25	20	16.0	24	1	LPER_STRFR
26	20	16.0	25	1	SCRK_LACIA
27	19.5	15.6	25	1	GPN2_JACOR
28	19	15.2	10	1	GON1_ALAMI
29	19	15.2	14	1	PPBL_PSEAE
30	19	15.2	16	1	CXAL_CONAL
31	19	15.2	18	1	AL13_CAKMA
32	19	15.2	19	1	ATPB_CANPA
33	19	15.2	19	1	COXR_THOOB

34	19	15.2	20	1	CD4_SHEEP	P05542	ovis aries
35	19	15.2	20	1	RNKD_PIG	P80551	sus scrofa
36	19	15.2	23	1	MLP_RANTE	P56924	rana tempor
37	19	15.2	24	1	ATPE_MICLU	P80286	micrococcus
38	19	15.2	24	1	VGL_BPAL3	P08766	bacterioph
39	19	15.2	25	1	COXO_ONCMY	P80334	oncorhynch
40	18	14.4	10	1	UPA4_HUMAN	P30090	homo sapien
41	18	14.4	12	1	UKA2_HUMAN	P31144	homo sapien
42	18	14.4	13	1	CXAZ_CONGE	P01520	canis geogr
43	18	14.4	13	1	ODPA_CANPA	P49823	canis faml
44	18	14.4	14	1	MASV_VESXA	P01515	vespa xanth
45	18	14.4	14	1	MT14_EISFO	P46979	eisenia foe

## ALIGNMENTS

RESULT 1	MISG_MISAN	STANDARD;	PRT;	21 AA.
ID	MISG_MISAN			
AC	P81474;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	MISGURIN.			
OS	Misgurnus anguillicaudatus.			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;			
OC	Cypriniformes; Cobitidae; Misgurnus.			
RN	[1]			
RF	SEQUENCE.			
RX	MEDLINE-97415401; PubMed-9271200;			
RA	Park C.B., Lee H.J., Park I.Y., Kim M.S., Kim S.C.;			
RT	"A novel antimicrobial peptide from the loach, Misgurnus			
RT	anguillicaudatus.";			
RL	FEBS Lett. 411:173-178(1997).			
CC	-I- FUNCTION: STRONG ANTIMICROBIAL ACTIVITY AGAINST SEVERAL			
CC	GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA AND FUNGI.			
CC	-I- MASS SPECTROMETRY: MW=2502; METHOD=MALDI.			
KW	Antibiotic; Fungicide.			
SQ	SEQUENCE 21 AA; 2502 MW; 4A6E9D0AB391BCF1 CRC64;			
QY	14 EHSFSLKKGAA 25			
Db	6 ELSKFS-KKGAA 16			
Query Match	22.8%; Score 28.5; DB 1; Length 21;			
Best Local Similarity	75.0%; Pred. No. 1.7e+02;			
Matches	9; Conservative 0; Mismatches 2; Indels 1; Gaps 1;			
RESULT 2	CR11_LITSP	STANDARD;	PRT;	25 AA.
ID	CR11_LITSP			
AC	P56226;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	CAERIN 1.1.			
OS	Litoria splendida, Litoria gilleni, and Litoria caerulea.			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;			
OC	Litoria.			
RN	[1]			
RP	SEQUENCE, AND MASS SPECTROMETRY OF 1.1; 1.1.1 AND 1.1.2.			
RC	SPECIES-L. SPLENDIDA; TISSUE-PAROTOID GLAND;			
RA	Stene D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;			
RT	"Peptides from Australian frogs. Structures of the caerins and			
RT	Caeridin 1 from Litoria splendida.";			
RL	J. Chem. Soc. Perkin Trans. 1:3173-3178(1992).			
[2]	SEQUENCE, AND MASS SPECTROMETRY OF 1.1; 1.1.1; 1.1.2 AND 1.1.4.			

CC SPECIES-L CAERINIA: TISSUE-PAROTOID GLAND;  
 RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;  
 "Peptides from Australian frogs. The structures of the caerins from  
 Litoria caerulea.";  
 RL J. Chem. Res. 138:910-936(1993).  
 RN [3]  
 RP SEQUENCE, AND MASS SPECTROMETRY OF 1.1; 1.1.1; 1.1.3; 1.1.5 TO 1.1.8.  
 RC SPECIES-L GILBERT: TISSUE-PAROTOID GLAND;  
 RA Waugh R.J., Stone D.J.M., Bowie J.H., Wallace J.C., Tyler M.J.;  
 "Peptides from Australian frogs. The structures of the caerins and  
 caeridins from Litoria gilberti.";  
 RL J. Chem. Res. 139:937-961(1993).  
 RN [4]  
 RP FUNCTION, AND STRUCTURE BY NMR.  
 RC SPECIES-L SPLENDIDA;  
 RX MEDLINE-97409981; PubMed-9266686;  
 RA Wong H., Bowie J.H., Carver J.A.;  
 "The solution structure and activity of caerin 1.1, an antimicrobial  
 peptide from the Australian green tree frog, Litoria splendida.";  
 RL Eur. J. Biochem. 247:545-557(1997).  
 CC -1- FUNCTION: ANTIBACTERIAL AND ANTIVIRAL PEPTIDES THAT ADOPT AN ALPHA  
 HELICAL CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH  
 CAERIN DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.  
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL  
 GLANDS.  
 CC -1- DOMAIN: CONTAINS TWO AMPHIPATHIC ALPHA HELIX REGIONS SEPARATED BY  
 A REGION OF LESS-DEFINED HELICITY AND GREATER FLEXIBILITY.  
 CC -1- PTM: THE MAJOR PRODUCT IS CAERIN 1.1; IN ADDITION, DIFFERENT  
 PEPTIDES ARE PRODUCED THAT ARE MISSING SOME AMINO ACID RESIDUES AT  
 THE N-TERMINUS OR C-TERMINUS. ALL ISOFORMS ARE NOT REPRESENTED IN  
 EACH SPECIES. CAERIN 1.1.1 AND CAERIN 1.1.4 ARE INACTIVE.  
 CC -1- MASS SPECTROMETRY: MM-2582; METHOD-FAB; RANGE-1-25.  
 CC -1- MASS SPECTROMETRY: MM-2412; METHOD-FAB; RANGE-3-25.  
 CC -1- MASS SPECTROMETRY: MM-2299; METHOD-FAB; RANGE-4-25.  
 CC -1- MASS SPECTROMETRY: MM-1421; METHOD-FAB; RANGE-13-25.  
 CC -1- MASS SPECTROMETRY: MM-2333; METHOD-FAB; RANGE-1-23.  
 CC -1- MASS SPECTROMETRY: MM-1626; METHOD-FAB; RANGE-1-16.  
 CC -1- MASS SPECTROMETRY: MM-1489; METHOD-FAB; RANGE-1-15.  
 CC -1- MASS SPECTROMETRY: MM-1180; METHOD-FAB; RANGE-1-12.  
 CC -1- MASS SPECTROMETRY: MM-915; METHOD-FAB; RANGE-1-10.  
 KW Antibiotic; Antiviral; Amphibian skin; Amidation.  
 FT PEPTIDE 1 25 CAERIN 1.1.  
 FT PEPTIDE 3 25 CAERIN 1.1.1.  
 FT PEPTIDE 4 25 CAERIN 1.1.2.  
 FT PEPTIDE 13 25 CAERIN 1.1.3.  
 FT PEPTIDE 1 23 CAERIN 1.1.4.  
 FT PEPTIDE 1 16 CAERIN 1.1.5.  
 FT PEPTIDE 1 15 CAERIN 1.1.6.  
 FT PEPTIDE 1 12 CAERIN 1.1.7.  
 FT PEPTIDE 1 10 CAERIN 1.1.8.  
 FT MOD\_RES 25 25 AMIDATION.  
 FT SEQUENCE 25 AA; 2585 MW; DBA5A460B0BE00 CRC64;  
 SO

Query Match 20.8%; Score 26; DB 1; Length 25;  
 Best Local Similarity 36.4%; Pred. No. 5.2e+02;  
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 5 LPRLAFTSEH 15  
 DB 14 LPHVYVIAEH 24

RESULT 3  
 CR12\_LITCE STANDARD; PRT; 25 AA.  
 AC P56227;  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE CAERIN 1.2.  
 OS Litoria caerulea.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;  
 CC Litoria.  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE-PAROTOID GLAND;  
 RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;  
 "Peptides from Australian frogs. The structures of the caerins from  
 Litoria caerulea.";  
 RL J. Chem. Res. 138:910-936(1993).  
 RN [1-]  
 RP FUNCTION: ANTIBACTERIAL PEPTIDE THAT ADOPTS AN ALPHA HELICAL  
 CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN  
 DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.  
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL  
 GLANDS.  
 CC -1- DOMAIN: CONTAINS TWO AMPHIPATHIC ALPHA HELIX REGIONS SEPARATED BY  
 A REGION OF LESS-DEFINED HELICITY AND GREATER FLEXIBILITY (BY  
 SIMILARITY).  
 CC -1- MASS SPECTROMETRY: MM-2552; METHOD-FAB.  
 KW Antibiotic; Amphibian skin; Amidation.  
 FT MOD\_RES 25 25 AMIDATION.  
 FT SEQUENCE 25 AA; 2555 MW; DBA5A460B1464C0 CRC64;  
 SO

Query Match 20.8%; Score 26; DB 1; Length 25;  
 Best Local Similarity 36.4%; Pred. No. 5.2e+02;  
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 5 LPRLAFTSEH 15  
 DB 14 LPHVYVIAEH 24

RESULT 4  
 CR13\_LITCE STANDARD; PRT; 25 AA.  
 AC P56228;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE CAERIN 1.3.  
 OS Litoria caerulea.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;  
 CC Litoria.  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE-PAROTOID GLAND;  
 RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;  
 "Peptides from Australian frogs. The structures of the caerins from  
 Litoria caerulea.";  
 RL J. Chem. Res. 138:910-936(1993).  
 RN [1-]  
 RP FUNCTION: ANTIBACTERIAL PEPTIDE THAT ADOPTS AN ALPHA HELICAL  
 CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN  
 DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.  
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL  
 GLANDS.  
 CC -1- DOMAIN: CONTAINS TWO AMPHIPATHIC ALPHA HELIX REGIONS SEPARATED BY  
 A REGION OF LESS-DEFINED HELICITY AND GREATER FLEXIBILITY (BY  
 SIMILARITY).  
 CC -1- MASS SPECTROMETRY: MM-2582; METHOD-FAB.  
 KW Antibiotic; Amphibian skin; Amidation.  
 FT MOD\_RES 25 25 AMIDATION.  
 FT SEQUENCE 25 AA; 2585 MW; DBA5A460B0E2F2 CRC64;  
 SO

Query Match 20.8%; Score 26; DB 1; Length 25;  
 Best Local Similarity 36.4%; Pred. No. 5.2e+02;  
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 5 LPRLAFTSEH 15  
 DB 14 LPHVYVIAEH 24

```

RESULT 5
ID CR14_LITGI STANDARD; PRT; 25 AA.
AC P56229;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CAERIN 1.4.
OS Litoria gilleni, and Litoria caerulea.
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY OF CAERIN 1.4.
RC SPECIES=L. CAERULEA; TISSUE-PAROTOID GLAND;
RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
RT "Peptides from Australian frogs. The structures of the caerins from
RT Litoria caerulea."
RL J. Chem. Res. 138:910-936(1993).
RN [2]
RP SEQUENCE, AND MASS SPECTROMETRY OF CAERIN 1.4.1.
RC SPECIES=L. GILLENII; TISSUE-PAROTOID GLAND;
RA Waugh R.J., Stone D.J.M., Bowie J.H., Wallace J.C., Tyler M.J.;
RT "Peptides from Australian frogs. The structures of the caerins and
RT caerins from Litoria gilleni."
RL J. Chem. Res. 139:937-961(1993).
CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE, THAT ADOPTS AN ALPHA HELICAL
CC CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN
CC DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL
CC GLANDS.
CC -1- DOMAIN: CONTAINS TWO AMPHIPATHIC ALPHA HELIX REGIONS SEPARATED BY
CC A REGION OF LESS-DEFINED HELICITY AND GREATER FLEXIBILITY (BY
CC SIMILARITY).
CC -1- MASS SPECTROMETRY: MM=2600; METHOD=FAB; RANGE=1-25.
CC -1- MASS SPECTROMETRY: MM=933; METHOD=FAB; RANGE=1-10.
KM Antibiotic; Amphibian skin; Amelation.
FT PEPTIDE 1 25 CAERIN 1.4.
CC 25 10 CAERIN 1.4.1.
FT MOD_RES 25 25 AMIDATION.
SQ SEQUENCE 25 AA; 2603 MW; D8A5BBA7DB80E00 CRC64;

Query Match 20.8%; Score 26; DB 1; Length 25;
Best Local Similarity 36.4%; Pred. No. 5.2e+02;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 LPRLIAFTSEH 15
DB 14 LPHVVPVIAEH 24

RESULT 6
ID GLUR_HUMAN STANDARD; PRT; 8 AA.
AC P07229;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE URINE GLYCOPROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE.
RA MEDLINE=72062338; Pubmed=5126885;
RA Lote C.J., Weiss J.B.;
RT "Identification in urine of a low-molecular-weight highly polar
RT glycopeptide containing cysteinyl-galactose."
RL Biochem. J. 123:25P-25P(1971).
CC -1- FUNCTION: THE IDENTITY OF THE GLYCOPROTEIN FROM WHICH THIS PEPTIDE
CC IS DERIVED IS UNKNOWN. NO PHYSIOLOGICAL FUNCTION HAS BEEN

```

```

CC ATTRIBUTED. AN ERYTHROCYTE MEMBRANE GLYCOPROTEIN HAVING A
CC SIMILAR STRUCTURE HAS ALSO BEEN FOUND.
DR PIR: A03188; XGHUEU.
KW Glycoprotein.
FT CARBOHYD 1 1 S-LINKED (GAL. . .)
SQ SEQUENCE 8 AA; 855 MW; C2D87A1F5B1EB1E CRC64;

Query Match 20.0%; Score 25; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.9e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 ESHS 17
DB 2 ESHS 5

RESULT 7
ID FRE4_LITIN STANDARD; PRT; 24 AA.
AC P82023;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE FREXATIN 4.
OS Litoria infraimata (giant tree frog).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=SKIN SECRETION;
RA MEDLINE=97368637; Pubmed=9225251;
RA Rattey M.J., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
RT "The structures of the frenatin peptides from the skin secretion of
RT the giant tree frog Litoria infraimata."
RL J. Pept. Sci. 2:117-124(1996).
CC -1- FUNCTION: WIDE SPECTRUM ANTIMICROBIAL PEPTIDE.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL
CC GLANDS.
CC -1- MASS SPECTROMETRY: MM=2493; METHOD=FAB.
KM Antibiotic; Amphibian skin.
SQ SEQUENCE 24 AA; 2495 MW; 6986D063947805A1 CRC64;

Query Match 20.0%; Score 25; DB 1; Length 24;
Best Local Similarity 83.3%; Pred. No. 7.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 20 LKGA 25
DB 6 LKGA 11

RESULT 8
ID CR15_LITCE STANDARD; PRT; 25 AA.
AC P56230;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CAERIN 1.5.
OS Litoria caerulea.
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-PAROTOID GLAND;
RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
RT "Peptides from Australian frogs. The structures of the caerins from
RT Litoria caerulea."
RL J. Chem. Res. 138:910-936(1993).

```

CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE, THAT ADOPTS AN ALPHA HELICAL  
 CC CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN  
 CC DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY  
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PANOTOID AND/OR ROSTRAL  
 CC GLANDS.  
 CC -1- DOMAIN: CONTAINS TWO AMPHIPATHIC ALPHA HELIX REGIONS SEPARATED BY  
 CC A REGION OF LESS-DEFINED HELICITY AND GREATER FLEXIBILITY (BY  
 CC SIMILARITY).  
 CC -1- MASS SPECTROMETRY: MW=2610; METHOD=FAB.  
 CC Antibiotic; Amphibian skin; Amidation.  
 FT MOD.RES 25  
 SQ SEQUENCE 25 AA: 2613 MW: 0FF5A464EA0BE12 CRC64;

Query Match 19.2%; Score 24; DB 1; Length 25;  
 Best Local Similarity 27.3%; Pred. No. 1.1e+03;  
 Matches 3; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 5 LPRLIAFTSEH 15  
 : : : :  
 DB 14 IPHVVPVIAEH 24

RESULT 9  
 VIP\_GADMO STANDARD; PRT; 25 AA.  
 AC P09684;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE VASOACTIVE INTESTINAL PEPTIDE (VIP) (FRAGMENT).  
 OS Gadus morhua (Atlantic cod).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 CC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadidae;  
 CC Gadus.  
 [1]  
 RP SEQUENCE.  
 RA Thwaites D.T., Young J., Thorndyke M.C., Dimaline R.;  
 RT "Isolation and characterisation of two teleost VIP's.";  
 RL Regul. Pept. 21:436-436(1988).  
 CC -1- FUNCTION: VIP CAUSES VASODILATION, LOWERS ARTERIAL BLOOD  
 CC PRESSURE, STIMULATES MYOCARDIAL CONTRACTILITY, INCREASES  
 CC GLYCOGENOLYSIS AND RELAXES THE SMOOTH MUSCLE OF TRACHEA, STOMACH  
 CC AND GALL BLADDER.  
 CC -1- MISCELLANEOUS: THERE IS ONLY A SINGLE DIFFERENCE BETWEEN THIS  
 CC SEQUENCE AND THAT OF CHICKEN (ALA-19 INSTEAD OF VAL-19).  
 CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.  
 DR PIR: J00361; J00361.  
 DR INTERPRO: IPR000532; .  
 DR PFMW: PF00123; hormone2; 1.  
 DR PROSITE: PS00260; GLUCAGON; 1.  
 KW Hormone.  
 FT NON\_TER 25  
 SQ SEQUENCE 25 AA: 2978 MW: 1573FF6F374DB7E4 CRC64;

Query Match 19.2%; Score 24; DB 1; Length 25;  
 Best Local Similarity 35.7%; Pred. No. 1.1e+03;  
 Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 11 FTSEHSHPSLKGA 24  
 : : : :  
 DB 6 FTDNYSRFRKQMAA 19

RESULT 10  
 CO4\_CAVPO STANDARD; PRT; 22 AA.  
 AC P19069;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)

DE COMPLEMENT C4 (FRAGMENT).  
 GN C4.  
 OS Cavia porcellus (Guinea pig).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Hystriognathii; Caviidae; Cavia.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE-80227885; PubMed-7391069;  
 RA Goldberger G., Abraham G.N., Williams J., Colten H.R.;  
 RT "NH2-terminal sequence analysis of pro-C4, the precursor of the  
 RT fourth component of guinea pig complement.";  
 RL J. Biol. Chem. 255:7071-7074(1980).  
 CC -1- FUNCTION: C4 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE  
 CC CLASSICAL PATHWAY OF THE COMPLEMENT SYSTEM. IT IS PROCESSED BY  
 CC ACTIVATED C1 WHICH REMOVE FROM THE ALPHA CHAIN THE C4A  
 CC ANAPHYLATOXIN  
 CC -1- SUBUNIT: THIS PROTEIN IS SYNTHESIZED AS A SINGLE-CHAIN PRECURSOR  
 CC AND, PRIOR TO SECRETION, IS ENZYMATICALLY CLEAVED TO FORM A TRIMER  
 CC OF NONIDENTICAL CHAINS (ALPHA, BETA, AND GAMMA).  
 DR PIR: A17267; A17267.  
 DR INTERPRO: IPR000020; .  
 DR INTERPRO: IPR001599; .  
 DR PROSITE: PS00477; ALPHA\_2\_MACROGLOBULIN; PARTIAL.  
 DR PROSITE: PS01177; ANAPHYLATOXIN\_1; PARTIAL.  
 DR PROSITE: PS01178; ANAPHYLATOXIN\_2; PARTIAL.  
 KW Complement pathway; Plasma; MHC III; Inflammatory response.  
 FT VARIANT 18  
 FT NON\_TER 22  
 FT L -> V.  
 SQ SEQUENCE 22 AA: 2452 MW: 2CD1B54D0B47A4 CRC64;

Query Match 18.4%; Score 23; DB 1; Length 22;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 PRLIAF 11  
 : : : :  
 DB 2 PRLLEF 7

RESULT 11  
 CH60\_THIPE STANDARD; PRT; 23 AA.  
 AC P29134;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN) (FRAGMENT).  
 GN MOPR OR GROEL.  
 OS Thiolobacillus ferrooxidans.  
 CC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN-ATCC 19859;  
 RX MEDLINE-93093401; PubMed-1360930;  
 RA Varela P., Jerez C.A.;  
 RT "Identification and characterization of GroEL and DnaK homologues in  
 RT Thiolobacillus ferrooxidans.";  
 RL FEMS Microbiol. Lett. 77:149-153(1992).  
 RN [2]  
 RP PHOSPHORYLATION.  
 RX MEDLINE-96242312; PubMed-9026439;  
 RA Seeger M., Osorio G., Jerez C.A.;  
 RT "Phosphorylation of GroEL, DnaK and other proteins from Thiolobacillus  
 RT ferrooxidans grown under different conditions.";  
 RL FEMS Microbiol. Lett. 138:129-134(1996).  
 CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND  
 CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS  
 CC CONDITIONS (BY SIMILARITY).  
 CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF  
 CC 7 SUBUNITS.  
 CC -1- INDUCTION: BY HEAT SHOCK.  
 CC -1- PTM: PHOSPHORYLATED ON THREONINE.

```

CC -1- MISCELLANEOUS: THIS PROTEIN SHOWS ATPASE ACTIVITY.
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
DR HSP60; P06139; 1A0N.
DR INTERPRO: IPR001844; -.
DR PROSITE: PS00296; CHAPERONINS_CPN60; PARTIAL.
KW Chaperone; ATP-binding; Heat shock; Phosphorylation.
FT NON_TER 23
SQ SEQUENCE 23 AA: 2536 MW; 2BB683872FB255D CRC64;

Query Match 18.0%; Score 22.5; DB 1; Length 23;
Best Local Similarity 33.3%; Pred. NO. 1.7e+03;
Matches 5; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

QY 9 IAFTEHSHFSLKKG 23
   |||:|:|:|:|
Db 5 VAF-AEHAREKMLRG 18

RESULT 12
LMT4_LOCM1
ID LMT4_LOCM1 STANDARD; PRT; 13 AA.
AC P41490;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE LOCUSTAMYOTROPIN 4 (LOW-MT 4).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pierisgata; Metazoa; Orthopteroidea; Orthoptera; Caellifera;
OC Acridomorpha; Acridoidea; Acrididae; Locusta.
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE-BRAIN;
RA Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., Kochansky J.P.,
RA de loof A.;
RT "Isolation, identification and synthesis of locustamytropin III and
RT IV, two additional neuropeptides of Locusta migratoria: members of the
RT locustamytropin peptide family.";
RL Insect Biochem. Mol. Biol. 22:447-452(1992).
CC -1- FUNCTION: POTENT MEDIATOR OF VISCERAL MUSCLE CONTRACTIONS ACTIVITY
CC (MYOTROPIC ACTIVITY). LOW-MT IV SEEMS TO BE A MORE POTENT MUSCLE
CC STIMULATOR THAN LOW-MT I, II AND III.
CC -1- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR INTERPRO: IPR001484; -.
DR PROSITE: PS00539; PYROKININ; 1.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 13
FT SEQUENCE 13 AA: 1553 MW; 20861943824D6698 CRC64;

Query Match 17.6%; Score 22; DB 1; Length 13;
Best Local Similarity 62.5%; Pred. NO. 1.1e+03;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GMAIPRL 8
   |||
Db 6 GMPSPRL 13

RESULT 13
CYCH_MOUSE
ID CYCH_MOUSE STANDARD; PRT; 18 AA.
AC Q61458;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CYCLIN H (FRAGMENT).
CC CNH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-CD-1; TISSUE-TESTIS;
RA Hall F.L., Wu L.;
RL Submitted (NOV-1994) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: MEMBER OF Cdk WHICH ACTIVATES CYCLIN-ASSOCIATED
CC CDK2/CDK4 BY THREONINE PHOSPHORYLATION. ITS EXPRESSION AND
CC ACTIVITY ARE CONSTANT THROUGHOUT THE CELL CYCLE. Cdk IS TIGHTLY
CC ASSOCIATED WITH A MULTIPROTEIN COMPLEX FTI1, WHICH PLAYS A DUAL
CC ROLE IN TRANSCRIPTION AND DNA REPAIR (BY SIMILARITY).
CC -1- SUBUNIT: MAMMALIAN Cdk CONTAINS THREE COMPONENTS: CDK7, CYCLIN H,
CC AND AN ASSEMBLY FACTOR CALLED MAT1.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN C SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X82441; CA57822.1; -.
DR HSP60; P51946; 1XU;
DR INTERPRO: IPR000553; -.
DR PROSITE: PS00292; CYCLINS; PARTIAL.
KW Cyclin; Cell cycle; Cell division; Nuclear protein;
KW Transcription regulation.
FT NON_TER 1
FT SEQUENCE 18 AA: 2105 MW; 92964DCF68B98C7 CRC64;

Query Match 17.6%; Score 22; DB 1; Length 18;
Best Local Similarity 57.1%; Pred. NO. 1.6e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 PRLIAFT 12
   |||
Db 10 PRLIMLT 16

RESULT 14
FAR7_ASCSU
ID FAR7_ASCSU STANDARD; PRT; 8 AA.
AC P43171;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE FMRFAMIDE-LIKE NEUROPEPTIDE AF7.
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascaridoidea;
OC Ascarididae; Ascaris.
RN [1]
RP SEQUENCE.
RX MEDLINE-95380362; PubMed-7651904;
RA Cowden C., Stretton A.O.W.;
RT "Eight novel FMRFamide-like neuropeptides isolated from the nematode
RT Ascaris suum.";
RL Peptides 16:491-500(1995).
CC -1- SIMILARITY: BELONGS TO THE FARF (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
CC KW Neuropeptide; Amidation.
FT MOD_RES 8
FT SEQUENCE 8 AA: 963 MW; 9CD40059D417687D CRC64;

Query Match 16.8%; Score 21; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. NO. 8.9e+04;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 PRLIAF 11
   |||

```

Db 3 PREFIRE 8

RESULT 15

LPRI\_ECOLI STANDARD: PRT: 16 AA.

AC P03058;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE HIS OPERON LEADER PEPTIDE (ATTENUATOR PEPTIDE).

GN HISL.

OS Escherichia coli, and Salmonella typhimurium.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

CC [1]

RP SEQUENCE FROM N.A.

RC SPECIES=E.COLI, AND S.TYPHIMURIUM; STRAIN=K12, AND LT2;

RX MEDLINE=89094829; PubMed=3062174;

RA Carlomagno M.S., Chiarloti L., Alfano P., Nappo A.G., Brunl C.B.;

RT "Structure and function of the Salmonella typhimurium and Escherichia coli K-12 histidine operons."

RL J. Mol. Biol. 203:585-606(1988).

RN [2]

RP SEQUENCE FROM N.A.

RC SPECIES=E.COLI;

RX MEDLINE=82059525; PubMed=6170941;

RA Verde P., Frunzio R., di Nocera P.P., Biasi F., Brunl C.B.;

RT "Identification, nucleotide sequence and expression of the regulatory region of the histidine operon of Escherichia coli K-12."

RL Nucleic Acids Res. 9:2075-2086(1981).

RN [3]

RP SEQUENCE FROM N.A.

RC SPECIES=E.COLI;

RX MEDLINE=79033821; PubMed=360215;

RA Dinocera P.P., Biasi F., Dilauro R., Frunzio R., Brunl C.B.;

RT "Nucleotide sequence of the attenuator region of the histidine operon of Escherichia coli K-12."

RL Proc. Natl. Acad. Sci. U.S.A. 75:4276-4280(1978).

RN [4]

RP SEQUENCE FROM N.A.

RC SPECIES=E.COLI; STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;

RA "The complete genome sequence of Escherichia coli K-12."

RL Science 277:1453-1474(1997).

RN [5]

RP SEQUENCE FROM N.A.

RC SPECIES=S.TYPHIMURIUM; STRAIN=LT2;

RX MEDLINE=79033822; PubMed=360216;

RA Barnes W.M.;

RT "DNA sequence from the histidine operon control region: seven histidine codons in a row."

RL Proc. Natl. Acad. Sci. U.S.A. 75:4281-4285(1978).

RN [6]

RP SEQUENCE FROM N.A.

RC SPECIES=S.TYPHIMURIUM; STRAIN=LT2;

RA Barnes W.M., Husson R.N., Whittier R.;

RL Submitted (AUG-1989) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: THIS PROTEIN IS INVOLVED IN THE ATTENUATION MECHANISM FOR THE CONTROL OF THE EXPRESSION OF THE HIS STRUCTURAL GENES.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: V00285; CAA23550.1; -

DR EMBL: V00284; CAA23548.1; -

DR EMBL: X13462; CAA31810.1; -

DR EMBL: V01371; CAA24656.1; -

DR EMBL: AE000293; AAC75079.1; -

DR EMBL: X13464; CAA31821.1; -

DR EMBL: J01804; AAA88613.1; -

DR PIR: A03594; LFECH.

DR ECOGENE: EG11269; HISL.

DR STYGENE: SG10166; HISL.

KW Histidine biosynthesis; leader peptide.

SQ SEQUENCE 16 AA; 2081 MW; C8B7907BF6283BBA CRC64;

Query Match 16.8%; Score 21; DB 1; Length 16;

Best Local Similarity 33.3%; Pred. No. 2e+03;

Matches 3; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 9 IAFTEHSH 17

Db 4 VQFKHHHH 12

RESULT 16

SETB\_SALTY STANDARD: PRT: 22 AA.

AC P33027;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE SUGAR EFFLUX TRANSPORTER B (FRAGMENT).

GN SETB.

OS Salmonella typhimurium.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; OC Salmonella.

CC [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89313694; PubMed=2546043;

RA Geerse R.H., Izzo F., Postma P.W.;

RT "The PEP: fructose phosphotransferase system in Salmonella typhimurium: PEP combines enzyme IIFru and pseudo-HPR activities."

RL Mol. Gen. Genet. 216:517-525(1989).

CC -1- FUNCTION: INVOLVED IN THE EFFLUX OF SUGARS. THE PHYSIOLOGICAL ROLE MAY BE THE DETOXIFICATION OF NON-METABOLIZABLE SUGAR ANALOGS. CAN TRANSPORT LACTOSE AND GLUCOSE (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE (PROBABLE).

CC -1- SIMILARITY: BELONGS TO THE SET FAMILY OF TRANSPORTERS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC DR EMBL: X14243; -; NOT\_ANNOTATED\_CDS.

DR STYGENE: SG10423; SETB.

KW Transport; Sugar transport; Transmembrane; Inner membrane.

FT TRANSMEM 13 >22 POTENTIAL.

FT NON\_TER 22 >22

SQ SEQUENCE 22 AA; 2291 MW; 2849C330172C5C2 CRC64;

Query Match 16.8%; Score 21; DB 1; Length 22;

Best Local Similarity 45.5%; Pred. No. 2.8e+03;

Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 AALPRLAFTS 13

Db 7 AAAPKSPDLTS 17

RESULT 17  
GAE6\_RANRU STANDARD; PRT; 24 AA.  
AC P80400;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DE 01-NOV-1997 (Rel. 35, Last annotation update)  
DE GAEGURIN-6.  
OS Rana rugosa (Frog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.  
RN [1]  
RP SEQUENCE.  
RC TISSUE-SKIN;  
RX MEDLINE-95091844; PubMed-7999137;  
RA Park J.M., Jung J.-E., Lee B.J.;  
RT "Antimicrobial peptides from the skin of a Korean frog, Rana rugosa".  
RT Biochem. Biophys. Res. Commun. 205:948-954(1994).  
CC -1- FUNCTION: HAS A NON-HEMOLYTIC ACTIVITY. HAS A BROAD SPECTRUM OF ACTIVITY AGAINST BOTH GRAM POSITIVE AND GRAM NEGATIVE BACTERIA, FUNGI AND PROTOZOAN.  
CC -1- SUBUNIT: MONOMER.  
CC -1- SUBCELLULAR LOCATION: SECRETED.  
CC -1- TISSUE SPECIFICITY: SKIN.  
CC -1- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN FAMILY.  
KM Amphibian skin; Antibiotic.  
FT DISULFID 18 24 BY SIMILARITY.  
SQ SEQUENCE 24 AA; 2610 MW; 09918123FF90CCFD CRC64;

Query Match 16.8%; Score 21; DB 1; Length 24;  
Best Local Similarity 54.5%; Pred. No. 3.1e+03;  
Matches 6; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 1 GMAA--LPRLI 9  
DB 7 GLAANFLPRLI 17

RESULT 18  
ALL1\_CYPDO STANDARD; PRT; 8 AA.  
AC P82152;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CYDIASTATIN 1.  
OS Cydia pomonella (Coddling moth).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.  
RN [1]  
RP SEQUENCE.  
RC TISSUE-LARVA;  
RX MEDLINE-98054539; PubMed-9392829;  
RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D., Davey M., East P.D., Thorpe A.;  
RT "Leptidopteran peptides of the allatostatin superfamily.";  
RL Peptides 18:1301-1309(1997).  
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
KM Neuropeptide; Amidation.  
FT MOD.RES 8 8 AMIDATION.  
SQ SEQUENCE 8 AA; 934 MW; C82879CA5B51F775 CRC64;

Query Match 16.0%; Score 20; DB 1; Length 8;  
Best Local Similarity 50.0%; Pred. No. 8.3e+04;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 13 SEHSHFSL 20

DB 1 SPHYNFGI 8  
RESULT 19  
GLEM\_HUMAN STANDARD; PRT; 10 AA.  
AC P02728;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DE 01-OCT-2000 (Rel. 40, Last annotation update)  
DE ERYTHROCYTE MEMBRANE GLYCOPOLYMER.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE-72034940; PubMed-5286858;  
RA Weiss J.B., Lote C.J., Bobinski H.;  
RT "New low molecular weight glycopeptide containing triglycosylcysteine in human erythrocyte membrane.";  
RL Nature New Biol. 234:25-26(1971).  
CC -1- PTM: S-LINKED GLYCAN CONSISTS OF GLC-GLC-GLC TRISACCHARIDE.  
CC -1- MISCELLANEOUS: THE IDENTITY OF THE GLYCOPROTEIN FROM WHICH THIS PEPTIDE IS DERIVED IS UNKNOWN. NO PHYSIOLOGICAL FUNCTION HAS BEEN ATTRIBUTED.  
CC PIR: A03187; XGHDE.  
KM Glycoprotein; Erythrocyte.  
FT CARBOHYD 1 1 S-LINKED (GLC...).  
SQ SEQUENCE 10 AA; 1049 MW; 239BFEEA1F5B1E8 CRC64;

Query Match 16.0%; Score 20; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 HSH 17  
DB 4 HSH 6

RESULT 20  
GONL\_PETMA STANDARD; PRT; 10 AA.  
AC P04378;  
DT 20-MAR-1987 (Rel. 04, Created)  
DT 20-MAR-1987 (Rel. 04, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE GONADOTROPIN I (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)  
(LULIBERIN I).  
OS Petromyzon marinus (Sea lamprey).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Hyperoartia;  
OC Petromyzontiformes; Petromyzontidae; Petromyzon.  
RN [1]  
RP SEQUENCE.  
RC TISSUE-BRAIN;  
RX MEDLINE-86168192; PubMed-3514603;  
RA Sherwood N.M., Sower S.A., Marshak D.R., Fraser B.A., Brownstein M.J.;  
RT "Primary structure of gonadotropin-releasing hormone from lamprey brain.";  
RL J. Biol. Chem. 261:4812-4819(1986).  
CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.  
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
CC PIR: A01412; RHLMGS.  
DR INTERPRO: IPR002012;  
DR PFAM: PF00446; GNRH; 1.  
DR PROSITE: PS00473; GNRH; 1.  
KM Hormone; Amidation; Hypothalamus.  
FT MOD.RES 1 1 PITRILIDONE CARBOXYLIC ACID.  
SQ SEQUENCE 10 AA; 1244 MW; 1E4B36237B1735AB CRC64;

Query Match 16.0%; Score 20; DB 1; Length 10;  
Best Local Similarity 60.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 17 HFSK 21  
1 1 1 1  
2 HYSLE 6

Db

RESULT 21  
MCA2\_RHOOP STANDARD; PRT; 15 AA.  
AC P56870;  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PUTATIVE MALEYLACTATE REDUCTASE II (EC 1.3.1.32) (FRAGMENT).  
OS Rhodococcus opacus.  
CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
CC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.  
RN [1]  
RP SEQUENCE.  
RC STRAIN=ICP;  
RX MEDLINE=98324954; PubMed=9657989;  
RA Seibert V., Kouratova E.M., Golovleva L.A., Schloemann M.;  
RT "Characterization of a maleylacetate reductase encoding region from  
Rhodococcus opacus ICP."  
RL J. Bacteriol. 180:3503-3508(1998).  
CC -1- PATHWAY: 3-CHLOROCATECHOL DEGRADATION (BETA-KETOADIPATE PATHWAY).  
CC THIS PATHWAY SERVES A VITAL ROLE IN THE BIODEGRADATION OF TOXIC  
CC AROMATIC COMPOUNDS INTRODUCED IN THE ENVIRONMENT BOTH AS NATURAL  
CC PRODUCTS AND AS INDUSTRIAL EFFLUENT.  
CC -1- SIMILARITY: BELONGS TO THE IRON-CONTAINING ALCOHOL DEHYDROGENASE  
CC FAMILY.  
CC INTERPRO: IPR001670. -  
DR PROSITE: PS00913; ADH\_IRON\_1; PARTIAL.  
DR PROSITE: PS00060; ADH\_IRON\_2; PARTIAL.  
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.  
FT NON\_TER 15  
SQ SEQUENCE 15 AA; 1884 MW; 58DA90DD038F025E CRC64;

Query Match 16.0%; Score 20; DB 1; Length 15;  
Best Local Similarity 57.1%; Pred. No. 2.8e+03;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 LPRLAF 11  
1 1 1 1  
8 LPRLIXF 14

Db

RESULT 22  
DFTS\_RAT STANDARD; PRT; 20 AA.  
ID DFTS\_RAT  
AC P07448;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 01-MAR-1989 (Rel. 10, Last annotation update)  
DE DENTINAL FLUID TRANSPORT-STIMULATING PEPTIDE (DFT-STIMULATING  
DE PEPTIDE).  
OS Rattus norvegicus (Rat).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=PAROTID GLAND;  
RX MEDLINE=87131231; PubMed=3815601;  
RA Yamamoto T., Kobayashi M., Kobayashi M., Yamamoto M., Nomura M.,  
RA Aonuma S.;  
RT "Isolation and amino acid sequence of dentinal fluid transport-  
RT stimulating peptide from rat parotid glands.";  
RL Chem. Pharm. Bull. 34:3803-3811(1986).

RN [2]  
RP CHARACTERIZATION.  
RX MEDLINE=67131708; PubMed=5297832;  
RA Steinman R.R.;  
RT "The movement of Acridavine hydrochloride through molars of rats on  
RT a cariogenic and non-cariogenic diet."  
RL J. South. Calif. Dent. Assoc. 35:151-157(1967).  
CC -1- FUNCTION: THIS PEPTIDE STIMULATES THE TRANSPORT OF DENTINAL FLUID,  
CC WHICH IS IMPORTANT FOR THE PREVENTION OF DENTAL CARIES.  
DR PIR: J00001; DIRT.  
DR Dental caries; Parotid gland; Hormone.  
SQ SEQUENCE 20 AA; 2165 MW; FA164F2B6AF80D5A CRC64;

Query Match 16.0%; Score 20; DB 1; Length 20;  
Best Local Similarity 27.8%; Pred. No. 3.7e+03;  
Matches 5; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 8 LIAPTSRHSFSLKKGAA 25  
1 1 1 1 1 1 1 1  
2 VIAMELOHNEPGRDSTA 19

Db

RESULT 23  
LPTR\_BACST STANDARD; PRT; 20 AA.  
ID LPTR\_BACST  
AC P05658;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE TETRACYCLINE RESISTANCE LEADER PEPTIDE.  
GN TEL.  
OS Bacillus stearothermophilus, Bacillus cereus, and  
OS Staphylococcus hyicus.  
OG Plasmid pTH15, Plasmid pBC16, and Plasmid pSTEL.  
CC Bacteria; Firmicutes; Bacillus/clostridium group;  
CC Bacillus/Staphylococcus group; Bacillus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC PLASMID=PTH15;  
RX MEDLINE=86031344; PubMed=2996983;  
RA Hoshino T., Ikeda T., Tomizuka N., Furukawa K.;  
RT "Nucleotide sequence of the tetracycline resistance gene of pTH15, a  
RT thermophilic Bacillus plasmid: comparison with staphylococcal TCR  
RT controls.";  
RL Gene 37:131-138(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC PLASMID=PRC16;  
RX MEDLINE=90221899; PubMed=2109312;  
RA Palva A., Vidgren G., Simonsen M., Rintala H., Laamanen P.;  
RT "Nucleotide sequence of the tetracycline resistance gene of pRC16  
RT from Bacillus cereus.";  
RL Nucleic Acids Res. 18:1635-1635(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECTIS-S-HYCUS; PLASMID=PSTEL;  
RX MEDLINE=92321725; PubMed=1622166;  
RA Schwarz S., Cardoso M., Wegener H.C.;  
RT "Nucleotide sequence and phylogeny of the tet(L) tetracycline  
RT resistance determinant encoded by plasmid pSTEL from Staphylococcus  
RT hyicus.";  
RL Antimicrob. Agents Chemother. 36:580-588(1992).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: D00006; BAA00004.1; -



```

DR EMBL: M11036; AAA22850.1; -
DR EMBL: X51366; CAA35750.1; -
DR EMBL: X60828; CAA43219.1; -
DR PIR: S09233; LEBSTU.
DR PIR: S23742; S23742.
KW Leader peptide; Antibiotic resistance; Plasmid.
SQ SEQUENCE 20 AA; 2253 MW; 1BD0F4A6CA231CA1 CRC64;

Query Match
Best Local Similarity 16.0%; Score 20; DB 1; Length 20;
Matches 4; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 13 SEHSFSLKGA 24
   | | | | |
Db 4 NECRVQLKEGS 15

RESULT 24
ID CYSF_TRIVA STANDARD; PRT; 22 AA.
AC P3404;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CYSTEINE PROTEINASE (EC 3.4.22.-) (FRAGMENT).
OS Trichomonas vaginalis.
OC Eukaryota; Parabasalida; Trichomonadida; Trichomonas.
RN [1]
RP SEQUENCE.
RX MEDLINE-93307628; PubMed-8319888;
RA Irvine J.W., Coombs G.H., North M.J.;
RT "Purification of cysteine proteinases from trichomonads using
bactracin-sepharose."
RL FEMS Microbiol. Lett. 110:113-120(1993).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
CC PAPAIN FAMILY OF THIOL PROTEASES.
DR MEROPS: C01.082; -
DR INTERPRO: IPR000169; -
DR PROSITE: PS00139; THIOL_PROTEASE_CYS; PARTIAL.
DR PROSITE: PS00639; THIOL_PROTEASE_HIS; PARTIAL.
DR PROSITE: PS00640; THIOL_PROTEASE_ASN; PARTIAL.
KM Hydrolase; Thiol protease.
FT UNSURE 1 1 OR K.
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2398 MW; 0EE40FD86661ACB CRC64;

Query Match
Best Local Similarity 16.0%; Score 20; DB 1; Length 22;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 KKGA 24
   | | | |
Db 9 KKGA 12

RESULT 25
ID LPER_STRFR STANDARD; PRT; 24 AA.
AC P45440;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE ERYTHROMYCIN RESISTANCE LEADER PEPTIDE (23S RNA METHYLASE LEADER
PEPTIDE).
OS Streptomyces fradiae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88169508; PubMed-3127381;
RA Kamimura S., Weisblum B.;

```

```

RT "Translational attenuation control of ermSF, an inducible resistance
RT determinant encoding RNA N-methyltransferase from Streptomyces
RT fradiae."
RL J. Bacteriol. 170:1800-1811(1988).
CC -1- FUNCTION: THIS PEPTIDE IS INVOLVED IN THE CONTROL MECHANISM OF
CC THE SYNTHESIS OF THE MACROLIDE-LINCOSAMIDE-STREPTOGRAMIN B
CC RESISTANCE PROTEIN.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-3 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M19269; AAA26741.1; -
KW Antibiotic resistance; Leader peptide.
SQ SEQUENCE 24 AA; 2529 MW; 756C1635F8C5F06A CRC64;

Query Match
Best Local Similarity 16.0%; Score 20; DB 1; Length 24;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GMAALP 6
   | | | |
Db 4 GMAALP 9

Search completed: February 5, 2001, 11:34:14
Job time: 356 sec

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 10:59:09 ; Search time 189.33 seconds  
(without alignments)  
3.612 Million cell updates/sec

Title: US-09-343-406-3

Perfect score: 98

Sequence: 1 ERGMIPSDLERLEAKQK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 141170

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_36.\*

```

1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	20	16 R88668	Human glutamic aci
2	93	94.9	20	16 R72281	Glutamic acid deca
3	93	94.9	20	21 Y59534	GAD65 fragment, pe
4	88	89.8	20	18 W18849	65 kD Glutamic aci
5	41	41.8	14	16 R76642	Peptide derived fr
6	41	41.8	14	18 W18850	65 kD Glutamic aci
7	36	36.7	22	20 Y19062	Lecithin:cholester
8	36	36.7	22	20 Y18808	Lecithin:cholester
9	36	36.7	22	20 Y18545	Lecithin:cholester
10	36	36.7	22	20 Y19316	Lecithin:cholester
11	34	34.7	20	20 Y25018	Streptokinase deri
12	32.5	33.2	22	20 Y19063	Lecithin:cholester

13	32.5	33.2	22	20 Y18809	Lecithin:cholester
14	32.5	33.2	22	20 Y18546	Lecithin:cholester
15	32.5	33.2	22	20 Y19317	Lecithin:cholester
16	32	32.7	16	14 R44062	Pulmonary surfacta
17	32	32.7	20	5 P40558	Sequence from eryt
18	32	32.7	21	14 R44058	Pulmonary surfacta
19	32	32.7	22	20 Y18973	Lecithin:cholester
20	32	32.7	22	20 Y18719	Lecithin:cholester
21	32	32.7	22	20 Y18456	Lecithin:cholester
22	32	32.7	22	20 Y19227	Lecithin:cholester
23	31.5	32.1	16	20 Y19185	Lecithin:cholester
24	31.5	32.1	16	20 Y18931	Lecithin:cholester
25	31.5	32.1	16	20 Y18672	Lecithin:cholester
26	31.5	32.1	16	20 Y19439	Lecithin:cholester
27	31	31.6	19	16 R71392	Plasamid pBMS-1 pol
28	30	30.6	12	18 W06561	Human hypoxia indu
29	30	30.6	15	20 Y19181	Lecithin:cholester
30	30	30.6	15	20 Y19184	Lecithin:cholester
31	30	30.6	15	20 Y18927	Lecithin:cholester
32	30	30.6	15	20 Y18930	Lecithin:cholester
33	30	30.6	15	20 Y18668	Lecithin:cholester
34	30	30.6	15	20 Y18671	Lecithin:cholester
35	30	30.6	15	20 Y19435	Lecithin:cholester
36	30	30.6	15	20 Y19438	Lecithin:cholester
37	30	30.6	18	20 Y19407	Lecithin:cholester
38	30	30.6	18	20 Y19153	Lecithin:cholester
39	30	30.6	18	20 Y18899	Lecithin:cholester
40	30	30.6	18	20 Y18636	Lecithin:cholester
41	30	30.6	19	20 W67988	Fragment of human
42	30	30.6	22	20 Y19348	Lecithin:cholester
43	30	30.6	22	20 Y19094	Lecithin:cholester
44	30	30.6	22	20 Y18840	Lecithin:cholester
45	30	30.6	22	20 Y18577	Lecithin:cholester

#### ALIGNMENTS

```

RESULT 1
R88668 ID R88668 standard; peptide; 20 AA.
XX AC R88668;
XX DT 05-MAR-1996 (first entry)
DE Human glutamic acid decarboxylase residues 306-325.
XX DE diabetes; T-cell subpopulation; detection; antigen production;
KW diagnosis; autoimmune disease.
XX OS Homo sapiens
XX PN DE4418091-A1.
XX PD 27-JUL-1995.
XX PF 24-MAY-1994; 94DE-4418091.
XX PR 04-FEB-1994; 94DE-4403522.
XX PR 20-JAN-1994; 94DE-4401629.
XX (ENDL/) ENDL J.
XX PA (BOEF) BOEHRINGER MANNHEIM GMBH.
XX PI Albert W, Dornmair K, Endl J, Jung G, Meinel E;
XX PI Stahl P, Schendel D;
XX DR WPI; 1995-264505/35.
XX PT Antigen-specific activated T-lymphocytes and their detection - by
XX PT interaction with inventive peptide(s) of peptide-MHC complexes;
XX PT useful in diagnosis of e.g. diabetes and autoimmune diseases

```

```

XX PS Claim 1; Page 14; 21pp; German.
XX CC R88668 is derived from residues 306-325 of human glutamic acid
XX CC decarboxylase and specifically reacts with a T-cell sub-population
XX CC isolated from recently diagnosed Type-1 diabetics. Pharmaceutical
XX CC compns. contg. this peptide and those shown in R88667 and R76642-62,
XX CC are useful for the diagnosis of a disease or predispositions of immune
XX CC system diseases, tumours, and autoimmune diseases, including diabetes.
XX CC The peptides are able to detect specific T-cell subpopulations that
XX CC are then used for antigen prodn., e.g. by reinjection.
XX SQ Sequence 20 AA;

Query Match 100.0%; Score 98; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.3e-10; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 ERGKMIPSDLERILEAKQ 20
Db 1 ergkmipsdlerrileakq 20

RESULT 2
R72281
ID R72281 standard; Peptide; 20 AA.
XX AC R72281;
XX DT 13-NOV-1995 (first entry)
XX DE Glutamic acid decarboxylase (GAD65) fragment.
XX KW Glutamic acid decarboxylase; GAD65; autoimmune disorders;
XX KW insulin-dependant diabetes mellitus; stiff man disease.
XX OS Homo sapiens.
XX PN W09507992-A.
XX PD 23-MAR-1995.
XX PF 24-AUG-1994; 94WO-US09478.
XX PR 17-SEP-1993; 93US-0123859.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;
XX DR WPT; 1995-131360/17.
XX PT New polypeptide fragments of glutamic acid decarboxylase - for
XX PT diagnosis and treatment of autoimmune disease, esp. insulin
XX PT dependent diabetes, also related nucleic acid, vectors,
XX PT antibodies, hybridoma(s) etc.
XX PS Claim 1; Page 76; 100pp; English.
XX CC Q86481 and Q86482 encode R71733 and R79105, rat and human glutamic
XX CC acid decarboxylase (GAD65) respectively, from which the GAD65
XX CC fragments described in R72261-R72298 were derived. These fragments
XX CC can be used to detect autoantibodies against GAD, e.g. to diagnose
XX CC and treat GAD-related autoimmune disorders, such as insulin
XX CC dependant diabetes mellitus or stiff man disease.
XX SQ Sequence 20 AA;

Query Match 94.9%; Score 93; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 3e-09; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

XX PS Claim 1; Page 14; 21pp; German.
XX CC R88668 is derived from residues 306-325 of human glutamic acid
XX CC decarboxylase and specifically reacts with a T-cell sub-population
XX CC isolated from recently diagnosed Type-1 diabetics. Pharmaceutical
XX CC compns. contg. this peptide and those shown in R88667 and R76642-62,
XX CC are useful for the diagnosis of a disease or predispositions of immune
XX CC system diseases, tumours, and autoimmune diseases, including diabetes.
XX CC The peptides are able to detect specific T-cell subpopulations that
XX CC are then used for antigen prodn., e.g. by reinjection.
XX SQ Sequence 20 AA;

Query Match 100.0%; Score 98; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.3e-10; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 ERGKMIPSDLERILEAKQ 20
Db 1 ergkmipsdlerrileakq 20

RESULT 2
R72281
ID R72281 standard; Peptide; 20 AA.
XX AC R72281;
XX DT 13-NOV-1995 (first entry)
XX DE Glutamic acid decarboxylase (GAD65) fragment.
XX KW Glutamic acid decarboxylase; GAD65; autoimmune disorders;
XX KW insulin-dependant diabetes mellitus; stiff man disease.
XX OS Homo sapiens.
XX PN W09507992-A.
XX PD 23-MAR-1995.
XX PF 24-AUG-1994; 94WO-US09478.
XX PR 17-SEP-1993; 93US-0123859.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Clare-Salzler MJ, Erlander MG, Kaufman DL, Tobin AJ;
XX DR WPT; 1995-131360/17.
XX PT New polypeptide fragments of glutamic acid decarboxylase - for
XX PT diagnosis and treatment of autoimmune disease, esp. insulin
XX PT dependent diabetes, also related nucleic acid, vectors,
XX PT antibodies, hybridoma(s) etc.
XX PS Claim 1; Page 76; 100pp; English.
XX CC Q86481 and Q86482 encode R71733 and R79105, rat and human glutamic
XX CC acid decarboxylase (GAD65) respectively, from which the GAD65
XX CC fragments described in R72261-R72298 were derived. These fragments
XX CC can be used to detect autoantibodies against GAD, e.g. to diagnose
XX CC and treat GAD-related autoimmune disorders, such as insulin
XX CC dependant diabetes mellitus or stiff man disease.
XX SQ Sequence 20 AA;

Query Match 94.9%; Score 93; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 3e-09; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

XX PS Claim 1; Page 14; 21pp; German.
XX CC R88668 is derived from residues 306-325 of human glutamic acid
XX CC decarboxylase and specifically reacts with a T-cell sub-population
XX CC isolated from recently diagnosed Type-1 diabetics. Pharmaceutical
XX CC compns. contg. this peptide and those shown in R88667 and R76642-62,
XX CC are useful for the diagnosis of a disease or predispositions of immune
XX CC system diseases, tumours, and autoimmune diseases, including diabetes.
XX CC The peptides are able to detect specific T-cell subpopulations that
XX CC are then used for antigen prodn., e.g. by reinjection.
XX SQ Sequence 20 AA;

Query Match 100.0%; Score 98; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.3e-10; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 ERGKMIPSDLERILEAKQ 19
Db 2 ergkmipsdlerrileakq 20

RESULT 3
Y59554
ID Y59554 standard; peptide; 20 AA.
XX AC Y59554;
XX DT 03-APR-2000 (first entry)
XX DE GAD65 fragment, peptide #21.
XX KW GAD65; glutamic acid decarboxylase 65; autoimmune disorder; IDDM;
XX KW insulin dependent diabetes mellitus; Stiff man disease; diagnosis;
XX KW therapy.
XX OS Homo sapiens.
XX PN US5998366-A.
XX PD 07-DEC-1999.
XX PF 09-APR-1997; 97US-0827618.
XX PR 07-JUN-1995; 95US-0485725.
XX PR 21-SEP-1990; 90US-0586536.
XX PR 18-JUN-1991; 91US-0716909.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Tobin AJ, Kaufman DL, Erlander MG;
XX DR WPI; 2000-095930/08.
XX PT Ameliorating glutamic acid decarboxylase associated autoimmune
XX PT disorders such as insulin dependent diabetes mellitus and Stiff man
XX PT disease -
XX PS Claim 1; Column 42; 61pp; English.
XX CC This sequence represents a fragment of the glutamic acid decarboxylase 65
XX CC (GAD65) protein. The invention relates to a method of ameliorating GAD
XX CC associated autoimmune disorders by administering a GAD65 peptide to the
XX CC patient. The method can be used for ameliorating GAD associated
XX CC autoimmune disorders such as IDDM (insulin dependent diabetes mellitus)
XX CC and Stiff man disease. GAD65 can also be useful for screening drugs that
XX CC alter GAD function, for generating monoclonal antibodies and in
XX CC immunoassays. GAD65 is an effective diagnostic tool for predicting IDDM
XX CC and the diagnosis is quite easy. It is also possible to obtain much
XX CC larger quantities of polypeptide via recombinant techniques than are
XX CC available from natural sources.
XX SQ Sequence 20 AA;

Query Match 94.9%; Score 93; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 3e-09; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

QY 1 ERGKMIPSDLERILEAKQ 19
Db 2 ergkmipsdlerrileakq 20

RESULT 4
W18849
ID W18849 standard; peptide; 20 AA.
XX AC W18849;

```

XX DT 05-JAN-1998 (first entry)  
 XX DE 65 kD Glutamic acid decarboxylase peptide fragment VIII.  
 XX KW GAD; 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;  
 KW insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;  
 KW prediagnosis; autoimmune; tumour; rheumatoid arthritis;  
 KW multiple sclerosis.  
 XX OS Synthetic.  
 XX PN DE19526561-A1.  
 XX PD 23-JAN-1997.  
 XX PF 20-JUL-1995; 95DE-1026561.  
 XX PR 20-JUL-1995; 95DE-1026561.  
 XX PA (BOEF ) BOEHRINGER MANNHEIM GMBH.  
 XX PI Donie F, Endl J, Ganz M, Jung G, Kientsch-engel R;  
 XX PI Pozzilli P, Stahl P;  
 XX DR WPI; 1997-088254/09.  
 XX PT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -  
 PT involving intradermal admin. of autoreactive substances  
 XX PS Claim 11; Page 10; 12pp; German.  
 XX CC W18842-70 are peptide fragments of the 65 kD human glutamic acid  
 CC decarboxylase (GAD). The fragments are autoreactive substances used for  
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is  
 CC determined by using a claimed method for diagnosis of cell-mediated  
 CC diseases or a predisposition to cell-mediated diseases, which is effected  
 CC by administering an autoreactive substance intradermally and establishing  
 CC the diagnosis on the basis of the occurrence or lack of a positive  
 CC reaction at the site of administration. The method is used for diagnosis  
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases  
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.  
 XX SQ Sequence 20 AA;

Query Match 89.8%; Score 88; DB 18; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 2.1e-08;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ERGKMIPDLERRILEAKOK 20  
 Db 1 ergkmipdslerileakqk 20  
 ||||| |||||  
 ||||| |||||

RESULT 5  
 R76642  
 ID R76642 standard; peptide; 14 AA.  
 XX AC R76642;  
 XX DT 05-MAR-1996 (first entry)  
 XX DE Peptide derived from human glutamic acid decarboxylase 1.  
 XX KW diabetes; T-cell subpopulation; detection; antigen production;  
 KW diagnosis; autoimmune disease.  
 XX OS Homo sapiens.  
 XX PN DE4418091-A1.  
 XX DR 27-JUL-1995.

XX PF 24-MAY-1994; 94DE-4418091.  
 XX PR 04-FEB-1994; 94DE-4403522.  
 XX PR 20-JAN-1994; 94DE-4401629.  
 XX PA (ENDL/) ENDL J.  
 XX PA (BOEF ) BOEHRINGER MANNHEIM GMBH.  
 XX PI Albert W, Dornmair K, Endl J, Jung G, Meinel E;  
 PI Stahl P, Schendel D;  
 XX DR WPI; 1995-264505/35.  
 XX PT Antigen-specific activated T-lymphocytes and their detection - by  
 PT interaction with inventive peptide(s) of peptide-MHC complexes;  
 PT useful in diagnosis of e.g. diabetes and autoimmune diseases  
 XX PS Claim 1; Fig 1; 21pp; German.  
 XX CC R76642-62 are derived from human glutamic acid decarboxylase and  
 CC specifically react with T-cell sub-populations isolated from recently  
 CC diagnosed Type-1 diabetics. Pharmaceutical compns. contg. these  
 CC peptides and those shown in R88668-72, are useful for the diagnosis of  
 CC a disease or predispositions of immune system diseases, tumours, and  
 CC autoimmune diseases, including diabetes. The peptides are able to detect  
 CC specific T-cell subpopulations that are then used for antigen prodn.,  
 CC e.g. by reinjection.  
 XX SQ Sequence 14 AA;

Query Match 41.8%; Score 41; DB 16; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGKMIPS 8  
 Db 7 ergkmips 14  
 |||||  
 |||||

RESULT 6  
 W18850  
 ID W18850 standard; peptide; 14 AA.  
 XX AC W18850;  
 XX DT 05-JAN-1998 (first entry)  
 XX DE 65 kD Glutamic acid decarboxylase peptide fragment 1.  
 XX KW GAD; 65 kD; human; glutamic acid decarboxylase; autoreactive; diagnosis;  
 KW insulin-dependent diabetes mellitus; IDDM; cell-mediated disease;  
 KW prediagnosis; autoimmune; tumour; rheumatoid arthritis;  
 KW multiple sclerosis.  
 XX OS Synthetic.  
 XX PN DE19526561-A1.  
 XX PD 23-JAN-1997.  
 XX PF 20-JUL-1995; 95DE-1026561.  
 XX PR 20-JUL-1995; 95DE-1026561.  
 XX PA (BOEF ) BOEHRINGER MANNHEIM GMBH.  
 XX PI Donie F, Endl J, Ganz M, Jung G, Kientsch-engel R;  
 PI Pozzilli P, Stahl P;  
 XX DR WPI; 1997-088254/09.

PT Skin test for diagnosis of cell-mediated diseases, esp. diabetes -  
 XX involving intradermal admin. of auto:reactive substances  
 PS Claim 11; Fig 1; 12pp; German.  
 XX  
 CC W18842-70 are peptide fragments of the 65 kD human glutamic acid  
 CC decarboxylase (GAD). The fragments are autoreactive substances used for  
 CC diagnosis of insulin-dependent diabetes mellitus (IDDM). The diagnosis is  
 CC determined by using a claimed method for diagnosis of cell-mediated  
 CC diseases or a predisposition to cell-mediated diseases, which is effected  
 CC by administering an autoreactive substance intradermally and establishing  
 CC the diagnosis on the basis of the occurrence or lack of a positive  
 CC reaction at the site of administration. The method is used for diagnosis  
 CC of autoimmune and tumour diseases, preferably T-cell-mediated diseases  
 CC such as rheumatoid arthritis, multiple sclerosis and especially IDDM.  
 XX  
 SQ Sequence 14 AA;

Query Match 41.8%; Score 41; DB 18; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ERGKWIPS 8  
 Db 7 ergkmips 14

RESULT 7  
 Y19062  
 ID Y19062 standard; Peptide; 22 AA.  
 XX  
 AC Y19062;  
 XX  
 DT 09-JUL-1999 (first entry)  
 XX  
 DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #129.

XX  
 KW Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;  
 KW human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;  
 KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;  
 KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;  
 KW septic shock.

XX Synthetic.  
 OS Homo sapiens.  
 OS  
 PN WO9916458-A1.  
 XX  
 PD 08-APR-1999.  
 XX  
 PF 28-SEP-1998; 98WO-US20326.  
 XX  
 PR 29-SEP-1997; 97US-0940096.

XX (BUTT/) BUTTNER K.  
 PA (CORN/) CORNUT I.  
 PA (DASS/) DASSEUX J.  
 PA (METZ/) METZ G.  
 PA (SEKU/) SEKUL R.

XX Buttner K, Cornut I, Dasseux J, Metz G, Sekul R;

XX WPI; 1999-277034/23.

XX Peptide agonists of apolipoprotein A-I

XX Example; Page 114; 254pp; English.

XX The present invention describes an agonist (A) of apolipoprotein A-I  
 CC (apoA-I) which is a 15-29 residue peptide, or analog, that forms an  
 CC amphipathic alpha-helix in presence of lipids. (A), and their lipid  
 CC complexes, are used to treat or prevent diseases associated with

CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,  
 CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I  
 CC deficiency; hypertriglyceridemia and metabolic syndrome, also for  
 CC treating septic shock. When labeled, (A) can also be used diagnostically  
 CC to measure serum levels of HDL, in particular the HDL subpopulation that  
 CC is involved in retrograde cholesterol transport, also to image HDL at  
 CC e.g. atherosclerotic streaks, and to raise antibodies. Y18934 to Y19187  
 CC represent lecithin:cholesterol acyltransferase (LCAT) activity  
 CC exhibiting core peptides, which are apoA-I agonists.

XX Sequence 22 AA;

Query Match 36.7%; Score 36; DB 20; Length 22;  
 Best Local Similarity 61.5%; Pred. No. 12;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 8 SDLERRILEAKOK 20  
 Db 7 ndllrellealqk 19

RESULT 8  
 Y18808  
 ID Y18808 standard; Peptide; 22 AA.  
 XX  
 AC Y18808;  
 XX  
 DT 09-JUL-1999 (first entry)  
 XX  
 DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #129.

XX  
 KW Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;  
 KW human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;  
 KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;  
 KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;  
 KW septic shock.

XX Synthetic.  
 OS Homo sapiens.  
 OS  
 PN WO9916408-A2.  
 XX  
 PD 08-APR-1999.

XX 28-SEP-1998; 98WO-US20328.  
 XX  
 PR 29-SEP-1997; 97US-0940093.

XX (BUTT/) BUTTNER K.  
 PA (CORN/) CORNUT I.  
 PA (DASS/) DASSEUX J.  
 PA (METZ/) METZ G.  
 PA (SEKU/) SEKUL R.

XX Buttner K, Cornut I, Dasseux J, Metz G, Sekul R;

XX WPI; 1999-277031/23.

XX Peptide agonists of apolipoprotein A-I

XX Example; Page 111; 152pp; English.

XX The present invention describes an agonist (A) of apolipoprotein A-I  
 CC (apoA-I) which is a 14-22 residue peptide, or analog, that forms an  
 CC amphipathic alpha-helix in presence of lipids. (A), and their lipid  
 CC complexes, are used to treat or prevent diseases associated with  
 CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,  
 CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I  
 CC deficiency; hypertriglyceridemia and metabolic syndrome, also for  
 CC treating septic shock. When labeled, (A) can also be used diagnostically  
 CC to measure serum levels of HDL, in particular the HDL subpopulation that  
 CC is involved in retrograde cholesterol transport, also to image HDL at

CC e.g. atherosclerotic streaks, and to raise antibodies. Y18680 to Y18933  
 CC represent lecithin:cholesterol acyltransferase (LCAT) activity  
 CC exhibiting core peptides, which are apoA-I agonists.

SQ Sequence 22 AA;

Query Match 36.7%; Score 36; DB 20; Length 22;  
 Best Local Similarity 61.5%; Pred. No. 12;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 8 SLEERRILEAKQK 20  
 Db 7 ndllrlelealqk 19

RESULT 9  
 Y18545  
 ID Y18545 standard; Peptide; 22 AA.

XX AC Y18545;

DT 09-JUL-1999 (first entry)

XX DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #129.  
 KW Gene therapy; apolipoprotein A-I; agonist; dyslipidemic disorder;  
 KW ApoA-I; cardiovascular disease; atherosclerosis; restenosis; LCAT;  
 KW hyperlipidemia; septic shock; lecithin:cholesterol acyltransferase.

XX OS Synthetic.

OS Homo sapiens.

XX WO9916409-A2.

PD 08-APR-1999.

PF 28-SEP-1998; 98WO-US20329.

PR 29-SEP-1997; 97US-0940136.

PA (BUTT/) BUTTNER K.

PA (CORN/) CORNUT I.

PA (DASS/) DASSEUX J.

PA (DUFO/) DUFOURCQ J.

PA (METZ/) METZ G.

PA (SEKU/) SEKUL R.

PI Buttner K, Cornut I, Dasseux J, Dufourcq J, Metz G;

PI Sekul R;

DR WPI; 1999-254921/21.

PT Nucleic acid encoding apolipoprotein A-I agonist peptides

PS Claim 18; Page 165; 232pp; English.

XX The present invention describes a nucleic acid (A) encoding an  
 CC apolipoprotein A-I (apoA-I) agonist (B) that is a peptide, or analog,  
 CC which forms an amphipathic alpha-helix in presence of lipids. (A),  
 CC optionally as a complex with lipids, and host cells that contain (A),  
 CC are useful for gene therapy, or prevention, of diseases associated with  
 CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,  
 CC atherosclerosis, restenosis, HDL (high density lipoprotein) and apoA-I  
 CC deficiency, hypertriglyceridemia and metabolic syndrome, also to treat  
 CC endotoxemia (septic shock). Host cells containing (A) can also be used  
 CC to study the role of apoA-I in lipid metabolism. (B) can be used  
 CC diagnostically, e.g. to measure serum HDL (particularly its  
 CC subpopulation involved in retrograde cholesterol transport) and for  
 CC imaging the circulatory system or HDL accumulations at fatty streaks.  
 CC The present sequence represents a peptide from the present invention.

XX Sequence 22 AA;

Query Match 36.7%; Score 36; DB 20; Length 22;  
 Best Local Similarity 61.5%; Pred. No. 12;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 8 SLEERRILEAKQK 20  
 Db 7 ndllrlelealqk 19

RESULT 10

ID Y19316  
 Y19316 standard; Peptide; 22 AA.

XX AC Y19316;

DT 14-JUL-1999 (first entry)

XX DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #129.  
 KW Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;  
 KW human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;  
 KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;  
 KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;  
 KW septic shock.

XX OS Synthetic.

OS Homo sapiens.

XX WO9916459-A1.

PD 08-APR-1999.

PF 28-SEP-1998; 98WO-US20327.

PR 29-SEP-1997; 97US-0940095.

PA (BUTT/) BUTTNER K.

PA (CORN/) CORNUT I.

PA (DASS/) DASSEUX J.

PA (DUFO/) DUFOURCQ J.

PA (METZ/) METZ G.

PA (SEKU/) SEKUL R.

PI Buttner K, Cornut I, Dasseux J, Dufourcq J, Metz G;

PI Sekul R;

DR WPI; 1999-277035/23.

PT Peptide agonists of apolipoprotein A-I

PS Claim 19; Page 156; 280pp; English.

XX The present invention describes an agonist (A) of apolipoprotein A-I  
 CC (apoA-I) which is a 15-29 residue peptide, or analog, that forms an  
 CC amphipathic alpha-helix in presence of lipids. (A), and their lipid  
 CC complexes, are used to treat or prevent diseases associated with  
 CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,  
 CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I  
 CC deficiency; hypertriglyceridemia and metabolic syndrome, also for  
 CC treating septic shock. When labeled, (A) can also be used diagnostically  
 CC to measure serum levels of HDL, in particular the HDL subpopulation that  
 CC is involved in retrograde cholesterol transport, also to image HDL at  
 CC e.g. atherosclerotic streaks, and to raise antibodies. Y19188 to Y19441  
 CC represent lecithin:cholesterol acyltransferase (LCAT) activity  
 CC exhibiting core peptides, which are apoA-I agonists.

XX Sequence 22 AA;

Query Match 36.7%; Score 36; DB 20; Length 22;  
 Best Local Similarity 61.5%; Pred. No. 12;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 8 SDLERRILEAKQK 20  
:|||:|||||  
Db 7 ndllrlealeqk 19

RESULT 11  
Y25018  
ID Y25018 standard; peptide; 20 AA.  
XX  
AC Y25018;  
XX  
DT 23-AUG-1999 (first entry)  
XX

Streptokinase derived peptide 10 for amellorization of cell death.

XX Streptokinase; cell death; amellorization; treatment; disease; aging; AIDS;  
XX cellular differentiation; physical insult; trauma; anoxia; hyperthermia;  
XX hypochemia; brain; spinal cord; kidney; heart; lung; liver; skin; HIV;  
XX viral disorder; hepatitis; retroviral; infection; encephalitis; FALS;  
XX Huntington's disease; cerebellar degeneration; cardiovascular disease;  
XX neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;  
XX familial amyotrophic lateral sclerosis; atherosclerosis; heart failure;  
XX infarction; heart disease; cardiomyopathy; hypertensive; myocardial;  
XX reperfusion injury; immune disease; rheumatoid arthritis; renal; retinal;  
XX systemic lupus erythematosus; insulin-dependent; diabetes mellitus;  
XX pernicious anaemia; dermatomyositis; enythemia nodosum; myasthenia gravis;  
XX Sjogren's syndrome; temporal arthritis; autoimmune polyarthritis;  
XX Wegener's granulomatosis; glomerulonephritis; anti-phospholipid syndrome;  
XX neoplastic disorder; leukemia; sarcoma; myeloma; carcinoma; neuroma;  
XX melanoma; cancer; breast; colon; cervix; prostate; Hodgkin's disease;  
XX non-Hodgkin's lymphoma; inflammatory disorder; stroke; ischemia;  
XX spinal cord; toxic insult; pulmonary; macular degeneration; cataract;  
XX pancreatitis; Crohn's disease; ulcerative colitis; motor neuron disease;  
XX Guillan Bare Syndrome; demyelinating disease; bypass surgery;  
XX chemotherapy; clozapine; AZI; anthracycline.

OS Synthetic.  
XX  
XX US5917013-A.  
XX  
XX 29-JUN-1999.  
XX  
XX 05-DEC-1996; 96US-0759599.  
XX  
XX 06-DEC-1995; 95US-0008233.  
XX  
XX 05-DEC-1996; 96US-0759599.  
XX  
XX (RABK/) RABKIN S W.  
XX  
XX Krystal G, Rabkin SW;  
PI  
XX WPI; 1999-394231/33.  
XX  
XX Peptides that amellorate cell death useful for treating conditions  
XX associated with cellular differentiation  
XX  
XX Example 3; Column 12; 15pp; English.

Y25009-Y25019 are novel peptides derived from streptokinase that  
amellorate cell death. The products of the invention and their encoding  
nucleic acids may be useful for treating diseases and conditions related  
to aging, cellular differentiation, physical insult (e.g. physical  
trauma, anoxia, hyperthermia, hypothermia, chemically induced damage, and  
trauma to the brain, spinal cord, kidney, heart, lungs, liver, skin and  
any other organ), viral disorders (e.g. hepatitis, retroviral infections,  
viral encephalitis, and AIDS/HIV), neurodegenerative disorders (e.g.  
Parkinson's disease, Alzheimer's disease, Huntington's disease,  
cerebellar degenerations, and familial amyotrophic lateral sclerosis  
(FALS)), cardiovascular disease (e.g. atherosclerosis, myocardial  
infarction, heart failure, cardiomyopathy, myocardial reperfusion injury,  
and hypertensive heart disease), immune disease (e.g. rheumatoid

CC arthritis, systemic lupus erythematosus, insulin-dependent, diabetes  
CC mellitus, lupus, pernicious anaemia, dermatomyositis, enythemia nodosum,  
CC Sjogren's syndrome, temporal arthritis, myasthenia gravis, Wegener's  
CC granulomatosis, glomerulonephritis, anti-phospholipid syndrome, and  
CC autoimmune polyarthritis), a neoplastic disorder (e.g. leukemia,  
CC sarcomas, myelomas, carcinomas, neuromas, melanoma, cancers of the  
CC breast, brain, colon, cervix, and prostate, Hodgkin's disease and  
CC non-Hodgkin's lymphoma), inflammatory disorders (e.g. inflammatory joint  
CC disorders and inflammatory induced cell damage to the eye, brain and  
CC other organs), ischemia or reperfusion injury (e.g. myocardial ischemia  
CC and reperfusion injury, renal ischemia, spinal cord ischemia and/or  
CC reperfusion injury, retinal ischemia or infarction, and stroke), toxic  
CC insult (e.g. liver toxicity, pulmonary toxicity, toxic damage to other  
CC organs from chemicals, radiation, and other noxious substances), macular  
CC degeneration, cataract formation, pancreatitis, Crohn's disease,  
CC ulcerative colitis, accelerated aging, spinal cord disease (e.g. motor  
CC neuron diseases, degeneration of the spinal cord, Guillan Bare Syndrome  
CC and demyelinating disease), bypass surgery, chemotherapy, AZT,  
CC chemically-induced reperfusion, and therapeutics such as clozapine, AZT,  
XX and anthracyclines.  
SQ Sequence 20 AA;

Query Match 34.7%; Score 34; DB 20; Length 20;  
Best Local Similarity 70.0%; Pred. No. 24;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 10 LERRILEAKQK 19  
|||:|||||  
Db 9 lerylleake 18

RESULT 12  
Y19063  
ID Y19063 standard; Peptide; 22 AA.  
XX  
AC Y19063;  
XX  
XX 09-JUL-1999 (first entry)  
XX  
XX Lecithin:cholesterol acyltransferase activation exhibiting peptide #130.  
XX  
XX Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;  
XX human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;  
XX cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;  
XX high density lipoprotein; hypertriglyceridemia; metabolic syndrome;  
XX septic shock.  
XX  
XX Synthetic.  
XX Homo sapiens.  
XX  
XX WO9916458-A1.  
XX  
XX 08-APR-1999.  
XX  
XX 28-SEP-1998; 98WO-US20326.  
XX  
XX 29-SEP-1997; 97US-0940096.  
XX  
XX (BUTT/) BUTTNER K.  
XX (CORN/) CORNUT I.  
XX (DASS/) DASSEUX J.  
XX (METZ/) METZ G.  
XX (SEKU/) SEKUL R.  
XX  
XX Buttner K, Cornut I, Dasseux J, Metz G, Sekul R;  
XX WPI; 1999-277034/23.  
XX  
XX Peptide agonists of apolipoprotein A-I  
XX  
XX Example; Page 114; 254pp; English.



XX The present invention describes an agonist (A) of apolipoprotein A-I  
 CC (apoA-I) which is a 15-29 residue peptide, or analog, that forms an  
 CC amphipathic alpha-helix in presence of lipids. (A), and their lipid  
 CC complexes, are used to treat or prevent diseases associated with  
 CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,  
 CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I  
 CC deficiency; hypertriglyceridemia and metabolic syndrome, also for  
 CC treating septic shock. When labeled, (A) can also be used diagnostically  
 CC to measure serum levels of HDL, in particular the HDL subpopulation that  
 CC is involved in retrograde cholesterol transport, also to image HDL at  
 CC e.g. atherosclerotic streaks, and to raise antibodies. Y18934 to Y19187  
 CC represent lecithin:cholesterol acyltransferase (LCAT) activity  
 CC exhibiting core peptides, which are apoA-I agonists.  
 XX Sequence 22 AA;

Query Match 33.2%; Score 32.5; DB 20; Length 22;  
 Best Local Similarity 64.3%; Pred. No. 47;  
 Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

OY 8 SDLERRILEA-KQK 20  
 Db 7 ndlllealkqk 20

RESULT 13  
 Y18809  
 ID Y18809 standard; Peptide; 22 AA.  
 XX AC Y18809;  
 XX 09-JUL-1999 (first entry)  
 DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #130.  
 KW Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;  
 KW human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;  
 KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;  
 KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;  
 KW septic shock.

XX Synthetic.  
 OS Homo sapiens.

XX WO9916408-A2.

XX 08-APR-1999.

XX 28-SEP-1998; 98WO-US20328.

XX 29-SEP-1997; 97US-0940093.

XX (BUTT/) BUTTNER K.

XX (CORN/) CORNUT I.

XX (DASS/) DASSEUX J.

XX (METZ/) METZ G.

XX (SEKU/) SEKUL R.

XX Buttner K, Cornut I, Dasseux J, Metz G, Sekul R;

XX WPI; 1999-277031/23.

XX Peptide agonists of apolipoprotein A-I

XX Example; Page 111; 152pp; English.

XX The present invention describes an agonist (A) of apolipoprotein A-I  
 CC (apoA-I) which is a 14-22 residue peptide, or analog, that forms an  
 CC amphipathic alpha-helix in presence of lipids. (A), and their lipid  
 CC complexes, are used to treat or prevent diseases associated with  
 CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease.

CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I  
 CC deficiency; hypertriglyceridemia and metabolic syndrome, also for  
 CC treating septic shock. When labeled, (A) can also be used diagnostically  
 CC to measure serum levels of HDL, in particular the HDL subpopulation that  
 CC is involved in retrograde cholesterol transport, also to image HDL at  
 CC e.g. atherosclerotic streaks, and to raise antibodies. Y18680 to Y18933  
 CC represent lecithin:cholesterol acyltransferase (LCAT) activity  
 CC exhibiting core peptides, which are apoA-I agonists.  
 XX Sequence 22 AA;

Query Match 33.2%; Score 32.5; DB 20; Length 22;  
 Best Local Similarity 64.3%; Pred. No. 47;  
 Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

OY 8 SDLERRILEA-KQK 20  
 Db 7 ndlllealkqk 20

RESULT 14  
 Y18546  
 ID Y18546 standard; Peptide; 22 AA.  
 XX AC Y18546;

XX 09-JUL-1999 (first entry)

XX DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #130.  
 KW Gene therapy; apolipoprotein A-I; agonist; dyslipidemic disorder;  
 KW ApoA-I; cardiovascular disease; atherosclerosis; restenosis; LCAT;  
 KW hyperlipidemia; septic shock; lecithin:cholesterol acyltransferase.  
 XX Synthetic.  
 OS Homo sapiens.

XX WO9916409-A2.

XX 08-APR-1999.

XX 28-SEP-1998; 98WO-US20329.

XX 29-SEP-1997; 97US-0940136.

XX (BUTT/) BUTTNER K.

XX (CORN/) CORNUT I.

XX (DASS/) DASSEUX J.

XX (DUFO/) DUFOURCQ J.

XX (METZ/) METZ G.

XX (SEKU/) SEKUL R.

XX Buttner K, Cornut I, Dasseux J, Dufourcq J, Metz G;

XX Sekul R;

XX WPI; 1999-254921/21.

XX Nucleic acid encoding apolipoprotein A-I agonist peptides

XX Claim 18; Page 166; 232pp; English.

XX The present invention describes a nucleic acid (A) encoding an  
 CC apolipoprotein A-I (apoA-I) agonist (B) that is a peptide, or analog,  
 CC which forms an amphipathic alpha-helix in presence of lipids. (A),  
 CC optionally as a complex with lipids, and host cells that contain (A),  
 CC are useful for gene therapy, or prevention, of diseases associated with  
 CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,  
 CC atherosclerosis, restenosis, HDL (high density lipoprotein) and apoA-I  
 CC deficiency; hypertriglyceridemia and metabolic syndrome, also to treat  
 CC endotoxemia (septic shock). Host cells containing (A) can also be used  
 CC to study the role of apoA-I in lipid metabolism. (B) can be used  
 CC diagnostically, e.g. to measure serum HDL (particularly its

CC subpopulation involved in retrograde cholesterol transport) and for  
 CC imaging the circulatory system or HDL accumulations at fatty streaks.  
 CC The present sequence represents a peptide from the present invention.

XX SQ Sequence 22 AA;

Query Match 33.2%; Score 32.5; DB 20; Length 22;  
 Best Local Similarity 64.3%; Pred. No. 47;  
 Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Oy 8 SDLERRILEA-KQK 20  
 Db 7 ndllrlealkqk 20

RESULT 15

Y19317  
 ID Y19317 standard; Peptide; 22 AA.

XX AC Y19317;

XX 14-JUL-1999 (first entry)

DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #130.

XX Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;  
 KW human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;  
 KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;  
 KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;  
 KW septic shock.

XX Synthetic.

OS Homo sapiens.

XX WO9916459-A1.

XX 08-APR-1999.

XX 28-SEP-1998; 98WO-US20327.

XX 29-SEP-1997; 97US-0940095.

XX (BUTT/) BUTTNER K.

PA (CORN/) CORNUT I.

PA (DASS/) DASSEUX J.

PA (DUFO/) DUFOURCQ J.

PA (METZ/) METZ G.

PA (SEKU/) SEKUL R.

PI Buttner K, Cornut I, Dasseux J, Dufourcq J, Metz G;

PI Sekul R;

XX WPI; 1999-277035/23.

XX Peptide agonists of apolipoprotein A-I

XX Claim 19; Page 156; 280pp; English.

XX The present invention describes an agonist (A) of apolipoprotein A-I  
 CC (apoA-I) which is a 15-29 residue peptide, or analog, that forms an  
 CC amphipathic alpha-helix in presence of lipids. (A), and their lipid  
 CC complexes, are used to treat or prevent diseases associated with  
 CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,  
 CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I  
 CC deficiency; hypertriglyceridemia and metabolic syndrome, also for  
 CC treating septic shock. When labeled, (A) can also be used diagnostically  
 CC to measure serum levels of HDL, in particular the HDL subpopulation that  
 CC is involved in retrograde cholesterol transport, also to image HDL at  
 CC e.g. atherosclerotic streaks, and to raise antibodies. Y19188 to Y19441  
 CC represent lecithin:cholesterol acyltransferase (LCAT) activity  
 CC exhibiting core peptides, which are apoA-I agonists.

XX

SQ Sequence 22 AA;

Query Match 33.2%; Score 32.5; DB 20; Length 22;  
 Best Local Similarity 64.3%; Pred. No. 47;  
 Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Oy 8 SDLERRILEA-KQK 20

Db 7 ndllrlealkqk 20

RESULT 16

R44062

ID R44062 standard; peptide; 16 AA.

XX AC R44062;

XX 14-MAY-1994 (first entry)

XX Pulmonary surfactant protein fragment.

KW SP-B; surfactant; phospholipids; pulmonary type II epithelial cells;  
 KW hyaline membrane disease; acute respiratory distress syndrome;  
 KW pulmonary drug delivery.

XX Synthetic.

OS WO9321221-A.

PN 28-OCT-1993.

XX 09-FEB-1993; 93WO-US01197.

XX 10-APR-1992; 92US-0866916.

XX (ABBO ) ABBOTT LAB.

XX Absolom DR, Baxter CH, Dhaon MK, Lundell EO, Sarin VK;

XX WPI; 1993-351656/44.

XX Compan. having surfactant activity when admixed with phospholipid  
 PT - comprising pulmonary surfactant protein fragments, used in  
 PT formulations for respiratory distress syndrome treatment  
 XX Claim 2; Page 29; 36pp; English.

XX The pulmonary surfactant protein is generated by cleavage of a  
 CC larger precursor protein synthesised by pulmonary type II epithelial  
 CC cells and affects the surfactant like activity of phospholipids.  
 CC The peptide shown is a C-terminal fragment (residues 63-78) which  
 CC can be used for the treatment of pulmonary surfactant deficient  
 CC diseases, including hyaline membrane disease, and/or abnormal states,  
 CC including acute respiratory distress syndrome and can also be used  
 CC for drug delivery systems.

XX See also R44056-70.

XX SQ Sequence 16 AA;

Query Match 32.7%; Score 32; DB 14; Length 16;  
 Best Local Similarity 40.0%; Pred. No. 40;  
 Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 3 GKMPDLSRRILEA 17

Db 1 grmlpqlvarlvira 15

RESULT 17

P40558

ID P40558 standard; protein; 20 AA.

XX P40558;  
 XX AC  
 XX DT 09-JAN-1992 (first entry)  
 XX DE Sequence from erythropoietin used to produce monoclonal antibodies  
 XX DE from cell line ATCC HB 8209.  
 XX KW Erythropoietin; hypoxia diagnosis; immunoassay.  
 XX PN WO8403152-A.  
 XX PD 16-AUG-1984.  
 XX PF 03-FEB-1984; 84WO-US00151.  
 XX PR 04-FEB-1983; 83US-0463724.  
 XX PA (AMGE-) AMGEN.  
 XX PA (KIRI-) KIRIN-AMGEN INC.  
 XX PA (KIRI ) KIRIN-AMGEN INC.  
 XX PI Egrie J;  
 XX PS Claim 11; Page 20; 25pp; English.  
 XX CC The monoclonal antibody is useful in the purification and isolation  
 XX CC of human erythropoietin from plasma by affinity chromatography. It  
 XX CC is also useful in immunoassays on human body fluids for  
 XX CC erythropoietin determination, e.g. in the diagnosis of hypoxia etc.  
 XX CC involving low or defective red blood cell prodn.  
 XX SQ Sequence 20 AA;  
 Query Match 32.7%; Score 32; DB 5; Length 20;  
 Best Local Similarity 77.8%; Pred. No. 51;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 10 LERRILEAK 18  
 Db III :IIII  
 12 lerylleak 20  
 RESULT 18  
 R44058  
 ID R44058 standard; peptide; 21 AA.  
 AC R44058;  
 XX DT 14-MAY-1994 (first entry)  
 XX DE Pulmonary surfactant protein fragment.  
 XX KW SP-B; surfactant; phospholipids; pulmonary type II epithelial cells;  
 KW hyaline membrane disease; acute respiratory distress syndrome;  
 KW pulmonary drug delivery.  
 XX OS Synthetic.  
 XX PN WO9321221-A.  
 XX PD 28-OCT-1993.  
 XX PF 09-FEB-1993; 93WO-US01197.  
 XX PR 10-APR-1992; 92US-0866916.

PA (ABBO ) ABBOTT LAB.  
 XX PI Absolom DR, Baxter CH, Dhaon MK, Lundell EO, Sarin VK;  
 XX DR WPI; 1993-351656/44.  
 XX PT Compsn. having surfactant activity when admixed with phospholipid  
 PT - comprising pulmonary surfactant protein fragments, used in  
 PT formulations for respiratory distress syndrome treatment  
 XX PS Claim 2; Page 29; 36pp; English.  
 XX CC The pulmonary surfactant protein is generated by cleavage of a  
 CC larger precursor protein synthesised by pulmonary type II epithelial  
 CC cells and affects the surfactant like activity of phospholipids.  
 CC The peptide shown is a C-terminal fragment (residues 58-78) which  
 CC can be used for the treatment of pulmonary surfactant deficient  
 CC diseases, including hyaline membrane disease, and/or abnormal states,  
 CC including acute respiratory distress syndrome and can also be used  
 CC for drug delivery systems.  
 XX CC See also R44056-70.  
 XX SQ Sequence 21 AA;  
 Query Match 32.7%; Score 32; DB 14; Length 21;  
 Best Local Similarity 40.0%; Pred. No. 54;  
 Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 GKXIPSDLERRILEA 17  
 Db I::I : I::I  
 6 grmlpqvlvarlvra 20  
 RESULT 19  
 Y18973  
 ID Y18973 standard; Peptide; 22 AA.  
 XX AC Y18973;  
 XX DT 09-JUL-1999 (first entry)  
 XX DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #40.  
 XX KW Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;  
 KW human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;  
 KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;  
 KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;  
 KW septic shock.  
 XX OS Synthetic.  
 OS Homo sapiens.  
 XX PN WO9916458-A1.  
 XX PD 08-APR-1999.  
 XX PF 28-SEP-1998; 98WO-US20326.  
 XX PR 29-SEP-1997; 97US-0940096.  
 XX PA (BUTT/) BUTTNER K.  
 PA (CORN/) CORNUT I.  
 PA (DASS/) DASSEUX J.  
 PA (METZ/) METZ G.  
 PA (SEKU/) SEKUL R.  
 XX PI Buttner K, Cornut I, Dasseux J, Metz G, Sekul R;  
 XX DR WPI; 1999-277034/23.  
 XX PT Peptide agonists of apolipoprotein A-I

PS Example; Page 109; 254pp; English.

XX The present invention describes an agonist (A) of apolipoprotein A-I

CC (apoA-I) which is a 15-29 residue peptide, or analog, that forms an

CC amphipathic alpha-helix in presence of lipids. (A), and their lipid

CC complexes, are used to treat or prevent diseases associated with

CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,

CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I

CC deficiency; hypertriglyceridemia and metabolic syndrome, also for

CC treating septic shock. When labeled, (A) can also be used diagnostically

CC to measure serum levels of HDL, in particular the HDL subpopulation that

CC is involved in retrograde cholesterol transport, also to image HDL at

CC e.g. atherosclerotic streaks, and to raise antibodies. Y18934 to Y19187

CC represent lecithin:cholesterol acyltransferase (LCAT) activity

CC exhibiting core peptides, which are apoA-I agonists.

XX Sequence 22 AA;

SQ

Query Match 32.7%; Score 32; DB 20; Length 22;

Best Local Similarity 53.8%; Pred. No. 57;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 8 SDLERRILEAKQK 20

Db :|||:||||

7 nellrellealqk 19

RESULT 20

Y18719

ID Y18719 standard; Peptide; 22 AA.

AC Y18719;

XX

DT 09-JUL-1999 (first entry)

XX

DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #40.

XX

XX Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;

KW human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;

KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;

KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;

KW septic shock.

XX

OS Synthetic.

OS Homo sapiens.

XX

PN WO9916408-A2.

XX

PD 08-APR-1999.

XX

PF 28-SEP-1998; 98WO-US20328.

XX

PR 29-SEP-1997; 97US-0940093.

XX

PA (BUTT/) BUTTNER K.

PA (CORN/) CORNUT I.

PA (DASS/) DASSEUX J.

PA (METZ/) METZ G.

PA (SEKU/) SEKUL R.

XX

PI Buttner K, Cornut I, Dasseux J, Metz G, Sekul R;

PI

XX WPI; 1999-277031/23.

DR

PT Peptide agonists of apolipoprotein A-I

XX

XX Example; Page 106; 152pp; English.

XX

XX The present invention describes an agonist (A) of apolipoprotein A-I

CC (apoA-I) which is a 14-22 residue peptide, or analog, that forms an

CC amphipathic alpha-helix in presence of lipids. (A), and their lipid

CC complexes, are used to treat or prevent diseases associated with

CC

CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,

CC atherosclerosis, restenosis, HDL (high density lipoprotein) and apoA-I

CC deficiency; hypertriglyceridemia and metabolic syndrome, also to treat

CC endotoxemia (septic shock). Host cells containing (A) can also be used

CC to study the role of apoA-I in lipid metabolism. (B) can be used

CC

CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,

CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I

CC deficiency; hypertriglyceridemia and metabolic syndrome, also for

CC treating septic shock. When labeled, (A) can also be used diagnostically

CC to measure serum levels of HDL, in particular the HDL subpopulation that

CC is involved in retrograde cholesterol transport, also to image HDL at

CC e.g. atherosclerotic streaks, and to raise antibodies. Y18934 to Y19187

CC represent lecithin:cholesterol acyltransferase (LCAT) activity

CC exhibiting core peptides, which are apoA-I agonists.

XX Sequence 22 AA;

SQ

Query Match 32.7%; Score 32; DB 20; Length 22;

Best Local Similarity 53.8%; Pred. No. 57;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 8 SDLERRILEAKQK 20

Db :|||:||||

7 nellrellealqk 19

RESULT 21

Y18456

ID Y18456 standard; Peptide; 22 AA.

AC Y18456;

XX

DT 09-JUL-1999 (first entry)

XX

DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #40.

XX

XX Gene therapy; apolipoprotein A-I; agonist; dyslipidemic disorder;

KW ApoA-I; cardiovascular disease; atherosclerosis; restenosis; LCAT;

KW hyperlipidemia; septic shock; lecithin:cholesterol acyltransferase.

XX

OS Synthetic.

OS Homo sapiens.

XX

PN WO9916409-A2.

XX

PD 08-APR-1999.

XX

PF 28-SEP-1998; 98WO-US20329.

XX

PR 29-SEP-1997; 97US-0940136.

XX

PA (BUTT/) BUTTNER K.

PA (CORN/) CORNUT I.

PA (DASS/) DASSEUX J.

PA (DUFO/) DUFOURCQ J.

PA (METZ/) METZ G.

PA (SEKU/) SEKUL R.

XX

PI Buttner K, Cornut I, Dasseux J, Dufourcq J, Metz G;

PI Sekul R;

XX

XX WPI; 1999-254921/21.

DR

PT Nucleic acid encoding apolipoprotein A-I agonist peptides

XX

XX Example; Page 138; 232pp; English.

XX

XX The present invention describes a nucleic acid (A) encoding an

CC apolipoprotein A-I (apoA-I) agonist (B) that is a peptide, or analog,

CC which forms an amphipathic alpha-helix in presence of lipids. (A),

CC optionally as a complex with lipids, and host cells that contain (A),

CC are useful for gene therapy, or prevention, of diseases associated with

CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,

CC atherosclerosis, restenosis, HDL (high density lipoprotein) and apoA-I

CC deficiency; hypertriglyceridemia and metabolic syndrome, also to treat

CC endotoxemia (septic shock). Host cells containing (B) can also be used

CC to study the role of apoA-I in lipid metabolism. (B) can be used

CC diagnostically, e.g. to measure serum HDL (particularly its  
CC subpopulation involved in retrograde cholesterol transport) and for  
CC imaging the circulatory system or HDL accumulations at fatty streaks.  
CC The present sequence represents a peptide from the present invention.

XX Sequence 22 AA;  
SQ

Query Match 32.7%; Score 32; DB 20; Length 22;  
Best Local Similarity 53.8%; Pred. No. 57;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 8 SLLRRILEAKQK 20  
Db :|||:|||||

RESULT 22  
Y19227  
ID Y19227 standard; Peptide; 22 AA.  
XX  
AC Y19227;  
XX  
DT 14-JUL-1999 (first entry)  
XX  
DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #40.  
XX  
KW Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;  
KW human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;  
KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;  
KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;  
KW septic shock.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
PN WO9916459-A1.  
XX  
PD 08-APR-1999.  
XX  
PF 28-SEP-1998; 98WO-US20327.  
XX  
PR 29-SEP-1997; 97US-0940095.  
XX  
PA (BUTT/) BUTTNER K.  
PA (CORN/) CORNUT I.  
PA (DASS/) DASSEUX J.  
PA (DUFO/) DUFORCQ J.  
PA (METZ/) METZ G.  
PA (SEKU/) SEKUL R.  
XX  
PI Buttner K, Cornut I, Dasseux J, Dufourcq J, Metz G;  
PI Sekul R;  
XX  
WPI; 1999-277035/23.  
XX  
PT Peptide agonists of apolipoprotein A-I  
XX  
PS Example; Page 118; 280pp; English.  
XX  
CC The present invention describes an agonist (A) of apolipoprotein A-I  
CC (apoA-I) which is a 15-29 residue peptide, or analog, that forms an  
CC amphipathic alpha-helix in presence of lipids. (A), and their lipid  
CC complexes, are used to treat or prevent diseases associated with  
CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,  
CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I  
CC deficiency; hypertriglyceridemia and metabolic syndrome, also for  
CC treating septic shock. When labeled, (A) can also be used diagnostically  
CC to measure serum levels of HDL, in particular the HDL subpopulation that  
CC is involved in retrograde cholesterol transport, also to image HDL at  
CC e.g. atherosclerotic streaks, and to raise antibodies. Y19188 to Y19441  
CC represent lecithin:cholesterol acyltransferase (LCAT) activity  
CC exhibiting core peptides, which are apoA-I agonists.

XX Sequence 22 AA;  
SQ

Query Match 32.7%; Score 32; DB 20; Length 22;  
Best Local Similarity 53.8%; Pred. No. 57;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 8 SLLRRILEAKQK 20  
Db :|||:|||||

RESULT 23  
Y19185  
ID Y19185 standard; Peptide; 16 AA.  
XX  
AC Y19185;  
XX  
DT 09-JUL-1999 (first entry)  
XX  
DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #256.  
XX  
KW Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;  
KW human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;  
KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;  
KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;  
KW septic shock.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
PN WO9916458-A1.  
XX  
PD 08-APR-1999.  
XX  
PF 28-SEP-1998; 98WO-US20326.  
XX  
PR 29-SEP-1997; 97US-0940096.  
XX  
PA (BUTT/) BUTTNER K.  
PA (CORN/) CORNUT I.  
PA (DASS/) DASSEUX J.  
PA (METZ/) METZ G.  
PA (SEKU/) SEKUL R.  
XX  
PI Buttner K, Cornut I, Dasseux J, Metz G, Sekul R;  
XX  
WPI; 1999-277034/23.  
XX  
PT Peptide agonists of apolipoprotein A-I  
XX  
PS Example; Page 121; 254pp; English.  
XX  
CC The present invention describes an agonist (A) of apolipoprotein A-I  
CC (apoA-I) which is a 15-29 residue peptide, or analog, that forms an  
CC amphipathic alpha-helix in presence of lipids. (A), and their lipid  
CC complexes, are used to treat or prevent diseases associated with  
CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,  
CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I  
CC deficiency; hypertriglyceridemia and metabolic syndrome, also for  
CC treating septic shock. When labeled, (A) can also be used diagnostically  
CC to measure serum levels of HDL, in particular the HDL subpopulation that  
CC is involved in retrograde cholesterol transport, also to image HDL at  
CC e.g. atherosclerotic streaks, and to raise antibodies. Y18934 to Y19187  
CC represent lecithin:cholesterol acyltransferase (LCAT) activity  
CC exhibiting core peptides, which are apoA-I agonists.

XX Sequence 16 AA;  
SQ

Query Match 32.1%; Score 31.5; DB 20; Length 16;  
Best Local Similarity 69.2%; Pred. No. 49;

Matches 9; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 9 DLERRILEA-KQK 20  
 ||| :||| |||  
 Db 4 dlfrllealkqk 16

RESULT 24  
 Y18931  
 ID Y18931 standard; Peptide; 16 AA.  
 AC Y18931;  
 XX  
 DT 09-JUL-1999 (first entry)  
 XX  
 DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #256.  
 KW Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;  
 KW human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;  
 KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;  
 KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;  
 KW septic shock.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9916408-A2.  
 XX  
 PD 08-APR-1999.  
 XX  
 PF 28-SEP-1998; 98WO-US20328.  
 XX  
 PR 29-SEP-1997; 97US-0940093.  
 XX  
 PA (BUTT/) BUTTNER K.  
 PA (CORN/) CORNUT I.  
 PA (DASS/) DASSEUX J.  
 PA (METZ/) METZ G.  
 PA (SEKU/) SEKUL R.  
 XX  
 PI Buttner K, Cornut I, Dasseux J, Metz G, Sekul R;  
 XX  
 WPI; 1999-277031/23.  
 XX  
 PT Peptide agonists of apolipoprotein A-I  
 XX  
 PS Example; Page 118; 152pp; English.  
 XX  
 CC The present invention describes an agonist (A) of apolipoprotein A-I  
 CC (apoA-I) which is a 14-22 residue peptide, or analog, that forms an  
 CC amphipathic alpha-helix in presence of lipids. (A), and their lipid  
 CC complexes, are used to treat or prevent diseases associated with  
 CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,  
 CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I  
 CC deficiency; hypertriglyceridemia and metabolic syndrome, also for  
 CC treating septic shock. When labeled, (A) can also be used diagnostically  
 CC to measure serum levels of HDL, in particular the HDL subpopulation that  
 CC is involved in retrograde cholesterol transport, also to image HDL at  
 CC e.g. atherosclerotic streaks, and to raise antibodies. Y18680 to Y18933  
 CC represent lecithin:cholesterol acyltransferase (LCAT) activity  
 CC exhibiting core peptides, which are apoA-I agonists.  
 XX  
 SQ Sequence 16 AA;

Query Match 32.1%; Score 31.5; DB 20; Length 16;  
 Best Local Similarity 69.2%; Pred. No. 49;  
 Matches 9; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 9 DLERRILEA-KQK 20  
 ||| :||| |||  
 Db 4 dlfrllealkqk 16

RESULT 25  
 Y18672  
 ID Y18672 standard; Peptide; 16 AA.  
 AC Y18672;  
 XX  
 DT 09-JUL-1999 (first entry)  
 XX  
 DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #256.  
 KW Gene therapy; apolipoprotein A-I; agonist; dyslipidemic disorder;  
 KW ApoA-I; cardiovascular disease; atherosclerosis; restenosis; LCAT;  
 KW hyperlipidemia; septic shock; lecithin:cholesterol acyltransferase.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9916409-A2.  
 XX  
 PD 08-APR-1999.  
 XX  
 PF 28-SEP-1998; 98WO-US20329.  
 XX  
 PR 29-SEP-1997; 97US-0940136.  
 XX  
 PA (BUTT/) BUTTNER K.  
 PA (CORN/) CORNUT I.  
 PA (DASS/) DASSEUX J.  
 PA (DUFO/) DUFOURCO J.  
 PA (METZ/) METZ G.  
 PA (SEKU/) SEKUL R.  
 XX  
 PI Buttner K, Cornut I, Dasseux J, Dufourcq J, Metz G;  
 XX  
 WPI; 1999-254921/21.  
 XX  
 PT Nucleic acid encoding apolipoprotein A-I agonist peptides  
 XX  
 PS Example; Page 202; 232pp; English.  
 XX  
 CC The present invention describes a nucleic acid (A) encoding an  
 CC apolipoprotein A-I (apoA-I) agonist (B) that is a peptide, or analog,  
 CC which forms an amphipathic alpha-helix in presence of lipids. (A),  
 CC optionally as a complex with lipids, and host cells that contain (A),  
 CC are useful for gene therapy, or prevention, of diseases associated with  
 CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,  
 CC atherosclerosis, restenosis, HDL (high density lipoprotein) and apoA-I  
 CC deficiency, hypertriglyceridemia and metabolic syndrome, also to treat  
 CC endotoxemia (septic shock). Host cells containing (B) can also be used  
 CC to study the role of apoA-I in lipid metabolism. (B) can be used  
 CC diagnostically, e.g. to measure serum HDL (particularly its  
 CC subpopulation involved in retrograde cholesterol transport) and for  
 CC imaging the circulatory system or HDL accumulations at fatty streaks.  
 CC The present sequence represents a peptide from the present invention.  
 XX  
 SQ Sequence 16 AA;

Query Match 32.1%; Score 31.5; DB 20; Length 16;  
 Best Local Similarity 69.2%; Pred. No. 49;  
 Matches 9; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 9 DLERRILEA-KQK 20  
 ||| :||| |||  
 Db 4 dlfrllealkqk 16

Search completed: February 5, 2001, 10:59:10  
 Job time: 633 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 11:27:23 ; Search time 265.72 Seconds  
(without alignments)

8.822 Million cell updates/sec

Title: US-09-343-406-3

Sequence: 1 ERGKIPSDLERILEAKQK 20

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 segs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 08

Maximum Match 100%  
Listing first 45 summaries

Database :

1: SP:REMBL.15:\*  
2: SP:archaea:\*  
3: SP:bacteria:\*  
4: SP:fungi:\*  
5: SP:human:\*  
6: SP:invertebrate:\*  
7: SP:mammal:\*  
8: SP:mhc:\*  
9: SP:organelle:\*  
10: SP:plant:\*  
11: SP:rodent:\*  
12: SP:virus:\*  
13: SP:vertebrate:\*  
14: SP:unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33.7	18	4	Q9POM3	Q9POM3 homo sapien
2	28.6	17	10	Q9SMC7	Q9SMC7 lycopersico
3	28.6	24	5	Q9U542	Q9U542 aedes aegypt
4	27.5	28.1	25	Q9TIG0	Q9TIG0 atelies belz
5	27.6	15	1	Q9UWH3	Q9UWH3 thermococu
6	27.6	24	5	Q24619	Q24619 dirosophila
7	27.6	25	11	Q26667	Q26667 mus musculu
8	26.5	11	11	P97330	P97330 mus musculu
9	26.5	17	2	Q46093	Q46093 campylobact
10	26.5	18	4	Q9UCN1	Q9UCN1 homo sapien
11	26.5	20	2	Q9R5F4	Q9R5F4 helicobacte
12	26.5	20	2	Q9R4W5	Q9R4W5 sus scrofa
13	25.5	15	6	Q9TRP0	Q9TRP0 sus scrofa
14	25.5	18	2	Q56610	Q56610 vibrio chol
15	25.5	18	4	Q9UCO2	Q9UCO2 homo sapien
16	25.5	21	12	Q87583	Q87583 chimpanzee
17	25.5	24	2	Q9R573	Q9R573 nitrosomona
18	25.5	25	5	Q26362	Q26362 eleutheria
19	24.5	20	6	Q9TR05	Q9TR05 oryctolagus

20	24	24.5	16	4	Q9UD21	Q9UD21 homo sapien
21	24	24.5	16	11	Q63039	Q63039 ratus norv
22	24	24.5	18	12	Q87587	Q87587 chimpanzee
23	24	24.5	20	2	Q9KJ72	Q9KJ72 mycoplasma
24	24	24.5	20	2	Q9KJ71	Q9KJ71 mycoplasma
25	24	24.5	20	2	Q9KJ70	Q9KJ70 mycoplasma
26	24	24.5	20	2	Q9KJ59	Q9KJ59 mycoplasma
27	24	24.5	20	2	Q9KJ58	Q9KJ58 mycoplasma
28	24	24.5	20	2	Q9KJ57	Q9KJ57 mycoplasma
29	24	24.5	20	2	Q9KJ56	Q9KJ56 mycoplasma
30	24	24.5	20	2	Q9KJ55	Q9KJ55 mycoplasma
31	24	24.5	20	2	Q9KJ54	Q9KJ54 mycoplasma
32	24	24.5	20	2	Q9KJ53	Q9KJ53 mycoplasma
33	24	24.5	20	2	Q9KJ52	Q9KJ52 mycoplasma
34	24	24.5	20	2	Q9KJ51	Q9KJ51 mycoplasma
35	24	24.5	20	2	Q9KJ50	Q9KJ50 mycoplasma
36	24	24.5	20	2	Q9KJ49	Q9KJ49 mycoplasma
37	24	24.5	20	2	Q9KJ48	Q9KJ48 mycoplasma
38	24	24.5	20	2	Q9KJ47	Q9KJ47 mycoplasma
39	24	24.5	20	2	Q9KJ46	Q9KJ46 mycoplasma
40	24	24.5	20	2	Q9KJ45	Q9KJ45 mycoplasma
41	24	24.5	20	2	Q9KJ44	Q9KJ44 mycoplasma
42	24	24.5	20	2	Q9KJ43	Q9KJ43 mycoplasma
43	24	24.5	20	2	Q9KJ42	Q9KJ42 mycoplasma
44	24	24.5	20	2	Q9KJ41	Q9KJ41 mycoplasma
45	24	24.5	20	2	Q9KJ40	Q9KJ40 mycoplasma

## ALIGNMENTS

RESULT 1  
Q9POM3 PRELIMINARY; PRT; 18 AA.  
Q9POM3  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE POLYOMYOSITIS/SCLEDERMA AUTOANTIGEN 2 (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bliskovsky V., Miller M., Mock B.;  
RT "Structure and localization of mouse PW-Sc1 genes."  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF152839; AAF73199.1; -;  
FT NON TER 18  
SQ SEQUENCE 18 AA; 2025 MW; AIF292BC77872D67 CRC64;

Query Match 33.7%; Score 33; DB 4; Length 18;  
Best Local Similarity 50.0%; Pred. No. 2.1e+02;  
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 7 PSDLERILEAKQK 20  
Db 4 PSTREPRVLASDOR 17  
RESULT 2  
Q9SMC7 PRELIMINARY; PRT; 17 AA.  
Q9SMC7  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE PUTATIVE PROLINE-RICH PROTEIN (FRAGMENT).  
CN CTDS.  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

CC Magnoliophyta: eudicotyledons: core eudicots: Asteridae: euasterids I:  
 CC Solanales: Solanaceae, Solanum.  
 CX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hoeberichts F.A., Orzaez D.V., Van der Plas L.H., Woltering E.J.;  
 RT "Changes in gene expression during Programmed Cell Death in tomato  
 cell suspensions."  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ230000; CAB61884.1; -.  
 FT NON\_TER  
 SQ SEQUENCE 17 AA; 1837 MW; E35DE1561000FFDC CRC64;

Query Match 28.6%; Score 28; DB 10; Length 17;  
 Best Local Similarity 55.6%; Pred. No. 1.1e+03;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GKMIPSDLE 11  
 DB 6 GKAPSDFO 14

RESULT 3  
 ID Q9U542 PRELIMINARY; PRT; 24 AA.  
 AC Q9U542;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE HISTONE (FRAGMENT).  
 OS Aedes aegypti (Yellowfever mosquito).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoides;  
 CC Culicidae; Aedes;  
 CX NCBI\_TaxID=7159;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA "tu 2".  
 RT "Genomic and evolutionary analysis of Peilal, a diverse family of  
 highly reiterated SINES in the yellow fever mosquito, Aedes aegypti."  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF107695; AAF20013.1; -.  
 FT NON\_TER  
 SQ SEQUENCE 24 AA; 2827 MW; 4E02A5F9AF1C950 CRC64;

Query Match 28.6%; Score 28; DB 5; Length 24;  
 Best Local Similarity 43.8%; Pred. No. 1.5e+03;  
 Matches 7; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 1 ERGKMIPSDLE--RRI 14  
 DB 4 KRVSIIPIRDIQTRRI 19

RESULT 4  
 ID Q9TGT0 PRELIMINARY; PRT; 25 AA.  
 AC Q9TGT0;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE ALPHA-CARDIAC ACTIN (FRAGMENT).  
 OS ACTC.  
 CC Ateles belzebuth chamek.  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles;  
 CX NCBI\_TaxID=118643;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Canavez F.C.;

RT "Gene mapping in Ateles paniscus chamek".  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF099178; AAF21852.1; -.  
 DR HSSP; P02568; IATN.  
 DR INTERPRO; IPR000279; -.  
 DR PFAM; PF00022; actin; 1.  
 FT NON\_TER  
 SQ SEQUENCE 25 AA; 2982 MW; 5242AF35E2169DCB CRC64;

Query Match 28.1%; Score 27.5; DB 6; Length 25;  
 Best Local Similarity 38.1%; Pred. No. 1.9e+03;  
 Matches 8; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 1 ERGKMIPSDLERIL-ERKOR 20  
 DB 5 ERGSEVYTAEREIVRDIKER 25

RESULT 5  
 ID Q9UWH3 PRELIMINARY; PRT; 15 AA.  
 AC Q9UWH3;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE KETO-VALINE-FERREDOXIN OXIDOREDUCTASE BETA-SUBUNIT, VOR BETA  
 DE (FRAGMENT).  
 OS Thermococcus litoralis.  
 CC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.  
 CX NCBI\_TaxID=2265;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=96146528; PubMed=8550513;  
 RA Helder J., Mai X., Adams M.W.;  
 RT "Characterization of 2-ketolysovalerate ferredoxin oxidoreductase, a  
 new and reversible coenzyme A-dependent enzyme involved in peptide  
 fermentation by hyperthermophilic archaea."  
 RL J. Bacteriol. 178:780-787(1996).  
 SQ SEQUENCE 15 AA; 1740 MW; 5C0D6371E91D4CAF CRC64;

Query Match 27.6%; Score 27; DB 1; Length 15;  
 Best Local Similarity 33.3%; Pred. No. 1.3e+03;  
 Matches 3; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 6 IPSDLERRI 14  
 DB 3 LPADVKKRL 11

RESULT 6  
 ID Q24619 PRELIMINARY; PRT; 24 AA.  
 AC Q24619;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE HSR-OMEGA DNA ENCODING HEAT SHOCK INDUCED RNA.  
 GN HSR-OMEGA.  
 OS Drosophila pseudoobscura (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 CX NCBI\_TaxID=7237;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89357452; PubMed=2475389;  
 RA Garbe J.C., Bendena W.G., Pardue M.L.;  
 RT "Sequence evolution of the Drosophila heat shock locus hsr omega. I.  
 The nonrepeated portion of the gene".  
 RL Genetics 122:403-415(1989).



DR EMBL: X16337; CAA34395.1; -  
 DR FLYBASE; FBgn0012703; Dpse\Hsr-omega.  
 KW Heat shock.  
 SQ SEQUENCE 24 AA; 2684 MW; 2EBBA19618BA5615 CRC64;

Query Match 27.6%; Score 27; DB 5; Length 24;  
 Best Local Similarity 44.4%; Pred. No. 2.1e+03;  
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 Oy 12 RRIELAKO 20  
 Db 6 KRIVNAKOR 14

RESULT 7  
 ID 026667 PRELIMINARY; PRT; 25 AA.  
 AC 026667;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE HOMEOBOX PROTEIN (FRAGMENT).  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92073357; PubMed=1720547;  
 RA Murtha M.T., Leckman J.F., Ruddle F.H.;  
 RT "Detection of homeobox genes in development and evolution."  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10711-10715(1991).  
 DR EMBL: M81666; AAA63528.1; -  
 DR INTERPRO: IPR001356; -  
 DR FRAM: PF000046; homeobox; 1.  
 KW Homeobox; DNA-binding; Nuclear protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 25 AA; 3056 MW; 4D03D8935520D30C CRC64;

Query Match 27.6%; Score 27; DB 11; Length 25;  
 Best Local Similarity 77.8%; Pred. No. 2.2e+03;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Oy 11 ERRILEAKO 19  
 Db 8 RRIELAKO 16

RESULT 8  
 ID P97330 PRELIMINARY; PRT; 11 AA.  
 AC P97330;  
 DT 01-MAY-1997 (TReMBLrel. 03, Created)  
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
 DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)  
 DE PROTEIN TYROSINE PHOSPHATASE, RECEPTOR TYPE, C POLYPEPTIDE-ASSOCIATED  
 DE PROTEIN (CD45-AP) (LSM-1).  
 GN PTPRCAP.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97124850; PubMed=8954783;  
 RA Brynys E., Mincheva A., Brynys R.M., Kirchessner H., Weltz S.,  
 RA Lichter P., Neuen S., Schraven B.;  
 RT "Sequence, genomic organization, and chromosomal localization of the  
 RT human LPAI (PTPRCAP) and mouse CD45-AP/LSM-1 genes."  
 RL Genomics 38:79-83(1996).  
 DR EMBL: X97268; CAA65923.1; -

DR MGI: 97811; Ptprcap.  
 SQ SEQUENCE 11 AA; 1150 MW; 50695413B5A772C7 CRC64;

Query Match 26.5%; Score 26; DB 11; Length 11;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Oy 5 MIPSDI 10  
 Db 4 LVPSDL 9

RESULT 9  
 ID 046093 PRELIMINARY; PRT; 17 AA.  
 AC 046093;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TReMBLrel. 08, Last annotation update)  
 DE HYPOTHETICAL PROTEIN (FRAGMENT).  
 OS Campylobacter jejuni.  
 CC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
 CC Campylobacter.  
 OX NCBI\_Taxid=197;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=81-176;  
 RX MEDLINE=95231295; PubMed=7715450;  
 RA Yao R., Burr D.H., Dolg P., Trust T.J., Niu H., Guerry P.;  
 RT "Isolation of motile and non-motile insertional mutants of  
 RT Campylobacter jejuni: the role of motility in adherence and invasion  
 RT of eukaryotic cells."  
 RL Mol. Microbiol. 14:883-893(1994).  
 DR EMBL: U09019; AAA68823.1; -  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 17 AA; 1989 MW; BFD25299957DB3B9 CRC64;

Query Match 26.5%; Score 26; DB 2; Length 17;  
 Best Local Similarity 46.2%; Pred. No. 2.1e+03;  
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 Oy 8 SDIERRILEAKO 20  
 Db 2 SDIANNILEDVKK 14

RESULT 10  
 ID 090CNI PRELIMINARY; PRT; 18 AA.  
 AC 090CNI;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
 DE (NA + K+)-ATPASE.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92375195; PubMed=1380674;  
 RA Chicz R.M., Urban R.G., Lane W.S., Gorga J.C., Stern L.J.,  
 RA Vignali D.A., Strominger J.L.;  
 RT "Predominant naturally processed peptides bound to HLA-DRI are derived  
 RT from MHC-related molecules and are heterogeneous in size."  
 RL Nature 358:764-768(1992).  
 DR INTERPRO: IPR001757; -  
 DR PFAM: PF00122; E1-E2\_ATPase; 1.  
 SQ SEQUENCE 18 AA; 1886 MW; 7F951335A9FA9070 CRC64;

Query Match 26.5%; Score 26; DB 4; Length 18;  
 Best Local Similarity 58.3%; Pred. No. 2.2e+03;  
 Matches 7; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

OY 6 IPSDLERRIL 17  
 1 1 1 1 1 1 1 1  
 DB 1 IPADL--RIISA 10

RESULT 11

OY9R5F4 PRELIMINARY; PRT; 20 AA.

AC OY9R5F4;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)  
 DE UREASE SMALL SUBUNIT (FRAGMENT).  
 OS Helicobacter mustelae.  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 RN NCBI\_TaxID=217;  
 RP SEQUENCE.  
 RX MEDLINE-93084378; PubMed-1452359;  
 RA Turbett G.R., Hoj P.B., Horne R., Mee B.J.;  
 RT "Purification and characterization of the urease enzymes of  
 Helicobacter species from humans and animals."  
 RL Infect. Immun. 60:5259-5266(1992).  
 DR INTERPRO: IPR002026; -;  
 DR PFAM: PF00547; urease\_gamma; 1.  
 SQ SEQUENCE 20 AA; 2275 MW; 29C9DE1D46D21435 CRC64;

Query Match 26.5%; Score 26; DB 2; Length 20;  
 Best Local Similarity 33.3%; Pred. No. 2.5e+03;  
 Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 4 KMIPSDLERRIL 15  
 1 1 1 1 1 1 1 1  
 DB 2 KLTPKELDKLML 13

RESULT 12

OY9R4W5 PRELIMINARY; PRT; 20 AA.

AC OY9R4W5;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)  
 DE 30 KDA MAJOR HEAT SHOCK PROTEIN (FRAGMENT).  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 RN NCBI\_TaxID=210;  
 RP SEQUENCE.  
 RX MEDLINE-95020803; PubMed-7935068;  
 RA Yokota K., Hirai Y., Haque M., Hayashi S., Isogai H., Sugiyama T.,  
 RA Nagamachi E., Tsukada Y., Fujii N., Oguma K.;  
 RT "Heat shock protein produced by Helicobacter pylori."  
 RL Microbiol. Immunol. 38:403-405(1994).  
 DR INTERPRO: IPR002026; -;  
 DR PFAM: PF00547; urease\_gamma; 1.  
 SQ SEQUENCE 20 AA; 2302 MW; 29C9DFBF6D21805 CRC64;

Query Match 26.5%; Score 26; DB 2; Length 20;  
 Best Local Similarity 33.3%; Pred. No. 2.5e+03;  
 Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 4 KMIPSDLERRIL 15  
 1 1 1 1 1 1 1 1

DB 2 KLTPKELDKLML 13

RESULT 13

ID OY9TRP0 PRELIMINARY; PRT; 15 AA.

AC OY9TRP0;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)  
 DE LOW MR ZONA PELLUCIDA BINDING PROTEIN (FRAGMENT).  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 RN NCBI\_TaxID=9823;  
 RP SEQUENCE.  
 RX MEDLINE-92378826; PubMed-1510840;  
 RA Parry R.V., Barker P.J., Jones R.;  
 RT "Characterization of low Mr zona pellucida binding proteins from boar  
 spermatozoa and seminal plasma."  
 RL Mol. Reprod. Dev. 33:108-115(1992).  
 SQ SEQUENCE 15 AA; 1525 MW; CC800E1028CF2058 CRC64;

Query Match 25.5%; Score 25; DB 6; Length 15;  
 Best Local Similarity 62.5%; Pred. No. 2.6e+03;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 ERGKMTPS 8  
 1 1 1 1 1 1 1 1  
 DB 4 KRGAATPS 11

RESULT 14

ID OY56610 PRELIMINARY; PRT; 18 AA.

AC OY56610;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, last annotation update)  
 DE ACQA (FRAGMENT).  
 GN ACQA.  
 OS Vibrrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.  
 RN NCBI\_TaxID=666;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C6706;  
 RX MEDLINE-97074686; PubMed-8917113;  
 RA Franco A., Peir-Eh Y., Johnson J., Barry E.M., Guerra H., Maurer R.,  
 RA Morris J.G.;  
 RT "Cloning and characterization of dnaE, encoding the catalytic subunit  
 of replicative DNA polymerase III, from Vibrrio cholerae strain  
 RT C6706."  
 RL Gene 175:281-283(1996).  
 DR EMBL: U30472; AAC44579.1; -;  
 DR NON\_TER 18  
 FT 18  
 SQ SEQUENCE 18 AA; 2153 MW; 18EBDAD212842EF CRC64;

Query Match 25.5%; Score 25; DB 2; Length 18;  
 Best Local Similarity 41.7%; Pred. No. 3.1e+03;  
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 9 DLERRILEAKOR 20  
 1 1 1 1 1 1 1 1  
 DB 7 DREKPIVELLETK 18

RESULT 15  
 ID OYUCQ2 PRELIMINARY; PRT; 18 AA.

AC Q9UCQ2;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE KU ANTIGEN 73 KDA PROTEIN FRACTION 9.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 NC NCB1\_Taxid-9606;  
 RN [1]  
 RP MEDLINE-92165807; PubMed-1537839;  
 RA Medychowski A., Henzel W., Huston L., Paslidis N., Ellerson D.,  
 RA McRae M., Seong D., Howard O.M., Deisseroth A.;  
 RT "Identification of proteins binding to Interferon-gamma  
 RT transcriptional enhancers in hematopoietic cells."  
 RL J. Biol. Chem. 267:4533-4540(1992).  
 SQ SEQUENCE 18 AA; 1909 MW; 369C9737383F87B9 CRC64;

Query Match 25.5%; Score 25; DB 4; Length 18;  
 Best Local Similarity 57.1%; Pred. No. 3.1e+03;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 GKMPD 9  
 DB 10 GLLPSD 16

RESULT 16  
 ID Q87583 PRELIMINARY; PRT; 21 AA.  
 AC Q87583;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE ISOLATE P051 REV (FRAGMENT).  
 GN REV.  
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
 OC Viruses; Retroviruses; Retroviridae; Lentivirus.  
 NC NCB1\_Taxid-11723;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-P051;  
 RX MEDLINE-97138325; PubMed-8985351;  
 RA Bibolet-Ruche F., Brenques C., Galat-Luong A., Galat G., Pourrut X.,  
 RA Vidal N., Veas F., Durand J.P., Cuny G.;  
 RT "Genetic diversity of simian immunodeficiency viruses from West  
 RT African green monkeys: evidence of multiple genotypes within  
 RT populations from the same geographical locale."  
 RL J. Virol. 71:307-313(1997).  
 DR EMBL: U37201; AAC56137.1; -  
 DR INTERPRO: IPR000625; -  
 DR PFAM: PF00424; REV; 1.  
 FT NON\_TER 21  
 SQ SEQUENCE 21 AA; 2599 MW; 6AD581D9CFE92914 CRC64;

Query Match 25.5%; Score 25; DB 12; Length 21;  
 Best Local Similarity 57.1%; Pred. No. 3.6e+03;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 7 PSDLRR 13  
 DB 5 PEELORR 11

RESULT 17  
 ID Q9R573 PRELIMINARY; PRT; 24 AA.  
 AC Q9R573;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE HYDROXYLAMINE OXIDOREDUCTASE (FRAGMENT).  
 OS Nitrosomonas europaea.  
 OC Bacteria; Proteobacteria; Beta subdivision;  
 OC Ammonia-oxidizing bacteria; Nitrosomonas.  
 NC NCB1\_Taxid-915;  
 RN [1]  
 RP MEDLINE-93315429; PubMed-8325841;  
 RA Arciero D.M., Hooper A.B.;  
 RT "Hydroxylamine oxidoreductase from Nitrosomonas europaea is a multimer  
 RT of an octa-heme subunit."  
 RL J. Biol. Chem. 268:14645-14654(1993).  
 DR HSSP: Q50925; 1FGJ.  
 SQ SEQUENCE 24 AA; 2663 MW; 7EC6694C5DCD9D64 CRC64;

Query Match 25.5%; Score 25; DB 2; Length 24;  
 Best Local Similarity 44.4%; Pred. No. 4.1e+03;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 ERGMPSD 9  
 DB 16 DRGATPKE 24

RESULT 18  
 ID Q26362 PRELIMINARY; PRT; 25 AA.  
 AC Q26362;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE AMP CLASS HOMEOBOX CNOX-1 (FRAGMENT).  
 OS Eleutheria dichotoma.  
 OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Anthomedusae;  
 OC Eleutheriidae; Eleutheria.  
 NC NCB1\_Taxid-13050;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-92078912; PubMed-1683896;  
 RA Schlierwater B., Murtha M., Dick M., Ruddle F.H., Buss L.W.;  
 RT "Homeoboxes in cnidarians."  
 RL J. Exp. Zool. 260:413-416(1991).  
 DR EMBL: S70691; AAB20572.1; -  
 FT HOMEOBOX: DNA-binding; Nuclear protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 25 AA; 3051 MW; 3BE2379355259EA3 CRC64;

Query Match 25.5%; Score 25; DB 5; Length 25;  
 Best Local Similarity 66.7%; Pred. No. 4.3e+03;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 11 ERRILEAKO 19  
 DB 8 ERRVLANQ 16

RESULT 19  
 ID Q9TRU5 PRELIMINARY; PRT; 20 AA.  
 AC Q9TRU5;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE FACTOR IX LIGHT CHAIN (GLA DOMAIN).  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 NC NCB1\_Taxid-9986;  
 RN [1]  
 RP SEQUENCE.

RX MEDLINE=92132840; PubMed=1776139;  
 RA Rao I.V., Hoang A.D.;  
 RT "Purification and characterization of rabbit factor IX and its  
 RL existence as a two-chain factor IX alpha in circulating plasma."  
 DR Thromb. Res. 64:57-68(1991).  
 DR HSSP; P00740; ICFH.  
 SQ SEQUENCE 20 AA; 2317 MW; 9368E94B4BF5800 CRC64;

Query Match 25.0%; Score 24.5; DB 6; Length 20;  
 Best Local Similarity 35.3%; Pred. No. 4.1e+03;  
 Matches 6; Conservative 4; Mismatches 4; Indels 3; Gaps 1;  
 OY 3 GKM--TPSDLERRITL 16  
 DB 4 GKLEFVSGNLERECIE 20

RESULT 20  
 ID Q9UD21 PRELIMINARY; PRT; 16 AA.  
 AC Q9UD21;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)  
 DE CYCLIN E-L (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95257942; PubMed=7739542;  
 RA Ohsudo M., Theodoras A.M., Schumacher J., Roberts J.M., Pagano M.;  
 RT "Human cyclin E, a nuclear protein essential for the G1-to-S phase  
 RL transition."  
 RL Mol. Cell. Biol. 15:2612-2624(1995).  
 SQ SEQUENCE 16 AA; 2089 MW; 777EFC69CA45E29C CRC64;

Query Match 24.5%; Score 24; DB 4; Length 16;  
 Best Local Similarity 50.0%; Pred. No. 3.8e+03;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 OY 11 ERRILEAKOK 20  
 DB 4 ERRERAKAKER 13

RESULT 21  
 ID Q63039 PRELIMINARY; PRT; 16 AA.  
 AC Q63039;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)  
 DE RAT TUMOR-ASSOCIATED ALDEHYDE DEHYDROGENASE (FRAGMENT).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 ON NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88209084; PubMed=3284529;  
 RA Harper K., Jones D.E. Jr., Brennan M.D., Lindahl R.;  
 RT "Characterization of a functional recombinant rat liver aldehyde  
 RL dehydrogenase: expression as a non-fusion protein in E. coli."  
 RL Biochem. Biophys. Res. Commun. 152:940-947(1988).  
 DR EMBL; M29320; AAA40722.1; -.  
 FT NON\_TER 16  
 SQ SEQUENCE 16 AA; 1836 MW; 7FFA5DC981665B83 CRC64;

Query Match 24.5%; Score 24; DB 11; Length 16;  
 Best Local Similarity 41.7%; Pred. No. 3.8e+03;  
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 8 SDLERRILEAKQ 19  
 DB 5 SDTVRAREARE 16

RESULT 22  
 ID Q87387 PRELIMINARY; PRT; 18 AA.  
 AC Q87387;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)  
 DE ISOLATE P055 REV (FRAGMENT).  
 GN REV.  
 OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 ON NCBI\_TaxID=11723;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=P055;  
 RX MEDLINE=97138325; PubMed=8985351;  
 RA Bibollet-Ruche F., Brengues C., Galat-Luong A., Galat G., Pourrut X.,  
 RA Vidal N., Veas F., Durand J.P., Cuny G.;  
 RT "Genetic diversity of simian immunodeficiency viruses from West  
 RT African green monkeys: evidence of multiple genotypes within  
 RL populations from the same geographical locale."  
 RL J. Virol. 71:307-313(1997).  
 DR EMBL; U37203; AAC56141.1; -.  
 DR INTERPRO: IPR000625; -.  
 DR PFAM: PF00424; REV. 1.  
 FT NON\_TER 18  
 SQ SEQUENCE 18 AA; 2189 MW; 59D92829141040F1 CRC64;

Query Match 24.5%; Score 24; DB 12; Length 18;  
 Best Local Similarity 50.0%; Pred. No. 4.3e+03;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 OY 7 PSDLEERRI 14  
 DB 5 PEELTRRL 12

RESULT 23  
 ID Q9KJ72 PRELIMINARY; PRT; 20 AA.  
 AC Q9KJ72;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
 DE FRUCTOSE BIPHOSPHATE ALDOLASE (FRAGMENT).  
 OS Mycoplasma mycoides mycoides LC.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 CC Entomoplasmataceae.  
 ON NCBI\_TaxID=44100;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=8794;  
 RX MEDLINE=20193983; PubMed=10727835;  
 RA Thlaucourt F., Lorenzon S., David A., Breard A.;  
 RT "Phylogeny of the Mycoplasma mycoides cluster as shown by sequencing  
 RL of a putative membrane protein gene."  
 RL Vet. Microbiol. 72:251-268(2000).  
 DR EMBL; AF162939; AAF6225.1; -.  
 FT NON\_TER 20  
 SQ SEQUENCE 20 AA; 2365 MW; 72542F75D7097AF5 CRC64;

Query Match 24.5%; Score 24; DB 2; Length 20;

Best Local Similarity 28.6%; Pred. No. 4.8e+03;  
Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 6 IPSDLERRILEAKO 19  
: : : : :  
DB 1 MPKLYHKKLYNAKO 14

DB 1 MPKLYHKKLYNAKO 14

Search completed: February 5, 2001, 11:27:24  
Job time: 1877 sec

## RESULT 24

O9KJT1 PRELIMINARY; PRT; 20 AA.  
AC O9KJT1;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, last annotation update)  
DE FRUCTOSE BIPHOSPHATE ALDOLASE (FRAGMENT).  
OS Mycoplasma mycoides mycoides LC.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;  
OC Entomoplasmataceae.  
OX NCBI\_TaxID=44100;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Y GOAT;  
RX MEDLINE-20193983; PubMed-10727835;  
RA Thlaucourt F., Lorenzon S., David A., Breard A.;  
RT "Phylogeny of the Mycoplasma mycoides cluster as shown by sequencing  
of a putative membrane protein gene."  
RL Vet. Microbiol. 72:251-268(2000).  
DR EMBL: AF162940; AAF26226.1; -.  
FT NON\_TER 20  
SQ SEQUENCE 20 AA; 2365 MW; 72542E75D7097AF5 CRC64;

Query Match 24.5%; Score 24; DB 2; Length 20;

Best Local Similarity 28.6%; Pred. No. 4.8e+03;  
Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 6 IPSDLERRILEAKO 19  
: : : : :  
DB 1 MPKLYHKKLYNAKO 14

## RESULT 25

O9KJT0 PRELIMINARY; PRT; 20 AA.  
AC O9KJT0;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, last annotation update)  
DE FRUCTOSE BIPHOSPHATE ALDOLASE (FRAGMENT).  
OS Mycoplasma mycoides mycoides LC.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;  
OC Entomoplasmataceae.  
OX NCBI\_TaxID=44100;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-266;  
RX MEDLINE-20193983; PubMed-10727835;  
RA Thlaucourt F., Lorenzon S., David A., Breard A.;  
RT "Phylogeny of the Mycoplasma mycoides cluster as shown by sequencing  
of a putative membrane protein gene."  
RL Vet. Microbiol. 72:251-268(2000).  
DR EMBL: AF162941; AAF26227.1; -.  
FT NON\_TER 20  
SQ SEQUENCE 20 AA; 2339 MW; 72542F6037097AF5 CRC64;

Query Match 24.5%; Score 24; DB 2; Length 20;

Best Local Similarity 28.6%; Pred. No. 4.8e+03;  
Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 6 IPSDLERRILEAKO 19  
: : : : :  
DB 1 MPKLYHKKLYNAKO 14

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 11:34:14 ; Search time 78.94 Seconds  
(without alignments)  
8.182 Million cell updates/sec

Title: US-09-343-406-3

Perfect score: 98  
Sequence: 1 ERGKNIPSDLERILEAKQK 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 1436

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31.6	25	1	ACP_ACICA	P80916 acinetobact
2	26.5	20	1	COX2_THUOB	P80974 thunnus obe
3	25.5	16	1	UVSK_BP76	Q06728 bacterioph
4	25.5	22	1	UVSK_BP72	Q06727 bacterioph
5	25.5	24	1	THD2_SALTY	P11954 salmonella
6	23.5	24.0	23	T2A_PARTE	Q27173 paramecium
7	23.5	24.0	23	T2B_PARTE	Q27174 paramecium
8	23.5	14	1	CAT2_FASHE	P80342 fasciola he
9	23.5	14	1	PPBL_PSEAE	P35482 pseudomonas
10	23.5	16	1	RL6_VIBPR	Q56715 vibrio prot
11	23.5	19	1	UKA1_HUMAN	P31940 homo sapien
12	23.5	20	1	COX2_ONCMY	P80329 oncorhynch
13	23.5	22	1	PA22_BOTSC	P80963 botriolechis
14	22.4	25	1	CYTB_THETS	P81064 thermomyces
15	22.4	23	1	CR42_LITCE	P56243 listeria cae
16	22.4	23	1	UDP_LACCA	F11962 lactobacill
17	22.4	16	1	PGK_VIBCH	P96154 vibrio chol
18	21.4	17	1	EFEG_THENO	Q01697 thermus aqu
19	21.4	18	1	FIXA_RHILE	P14313 rhizobium 1
20	21.4	23	1	VG22_BP76	P21597 bacterioph
21	21.4	24	1	AMAA_BACTR	P37356 bacillus th
22	21.4	25	1	RL41_AGABI	F78589 agarticus bl
23	20.4	6	1	ACPH_RABIT	P25154 oryctolagus
24	20.4	14	1	YMPX_XANCP	Q03397 xanthomonas
25	20.4	15	1	UE15_HORVU	P34938 hordeum vul
26	20.4	17	1	RS13_PARDE	P72180 paracoccus
27	20.4	19	1	ADG_CLOPA	P81336 clostridium
28	20.4	19	1	MIFR_TRISP	P81339 trichinella
29	20.4	20	1	AMP_FUSNU	F81307 fusobacteri
30	20.4	21	1	PEDB_HYDAT	P80577 hydatra atten
31	20.4	21	1	RL41_YEAST	P05746 saccharomyc
32	19.4	9	1	ULAK_MOUSE	P99031 mus musculi
33	19.4	13	1	IDHA_CANFA	P54836 canis famli

34	19	19.4	13	1	UP51_UPRIN	P82036 uperoleia 1
35	19	19.4	19	1	FI8B_VULVU	P14482 vulpes vulp
36	19	19.4	23	1	NIFD_ANASL	P33177 anasdoma sp
37	19	19.4	24	1	LEC_CROST	P16351 crocotalaria
38	19	19.4	24	1	RS5_VIBPR	P52856 vibrio prot
39	19	19.4	24	1	WORC_MERTM	P80909 methanobact
40	19	19.4	25	1	RL41_QUEBU	Q82713 quercus sub
41	18.5	18.9	23	1	COXK_CANFA	Q9728 canis famli
42	18	18.4	10	1	GAJU_HUMAN	P01358 homo sapien
43	18	18.4	10	1	URE3_MORMO	P17339 morganelia
44	18	18.4	14	1	YGDH_THRAC	Q05213 thermoplasm
45	18	18.4	15	1	DCM1_PSECA	P19920 pseudomonas

## ALIGNMENTS

RESULT	1	ALIGNMENTS
ACP_ACICA	STANDARD;	PRT; 25 AA.
AC	P80916;	
DT	01-NOV-1997 (Rel. 35, Created)	
DT	01-NOV-1997 (Rel. 35, Last sequence update)	
DT	01-NOV-1997 (Rel. 35, Last annotation update)	
DE	ACTL CARRIER PROTEIN (ACP) (FRAGMENT).	
GN	ACP.	
OS	Acinetobacter calcoaceticus.	
OC	Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;	
OC	Acinetobacter.	
RN	[1]	
RP	SEQUENCE.	
RX	STRAIN-ATCC 33604;	
RX	MEDLINE-97315245; PubMed-9171419;	
RA	Tang L., Weisborn A.C., Kennedy E.P.;	
RT	"Domains of Escherichia coli acyl carrier protein important for	
RT	membrane-derived-oligosaccharide biosynthesis.";	
RL	J. Bacteriol. 179:3697-3705(1997).	
CC	-1- FUNCTION: THIS PROTEIN IS THE CARRIER OF THE GROWING FATTY ACID	
CC	CHAIN IN FATTY ACID BIOSYNTHESIS.	
CC	-1- PATHWAY: KEY COMPONENT IN DE NOVO FATTY ACID BIOSYNTHESIS.	
CC	-1- PTM: THE GROWING FATTY ACID CHAINS ARE COVALENTLY BOUND TO THE	
CC	4'-PHOSPHOPANTHETHEINE PROSTHETIC GROUP.	
DR	INTERPRO: IPR000255.	
DR	PROSITE: PS00012; PHOSPHOPANTHETHEINE, PARTIAL.	
DR	PROSITE: PS00075; ACP_DOMAIN; PARTIAL.	
FW	Fatty acid biosynthesis; Phosphopantetheine.	
FT	NON_TER	
FT	25	
FT	25	
SEQ	SEQUENCE 25 AA: 2886 MW: 38186ACIC7A3907 CRC64;	
Query Match	31.6%; Score 31; DB 1; Length 25;	
Best Local Similarity	60.0%; Pred. No. 1.1e+02;	
Matches	6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;	
QY	8 SFLERILEA 17	
QY	11:11:11:1	
DB	1 SDIEORIKQA 10	
COXB_THUOB	STANDARD;	PRT; 20 AA.
AC	P80974;	
DT	01-NOV-1997 (Rel. 35, Created)	
DT	01-NOV-1997 (Rel. 35, Last sequence update)	
DT	15-JUL-1999 (Rel. 38, Last annotation update)	
DE	CYTOCHROME C OXIDASE POLYPEPTIDE VB (EC 1.9.3.1) (FRAGMENT).	
OS	Thunnus obesus (Bigeye tuna).	
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;	
OC	Acanthomorpha; Acanthopterygii; Perciformes; Scombroidei;	
OC	Scombridae; Thunnus.	
RN	[1]	





```

RX MEDLINE-83023208; PubMed-6751404;
RA Kim S.S., Data P.;
RT "Chemical characterization of biodegradative threonine dehydratases
RL from two enteric bacteria.";
RL Blochm. Biophys. Acta 706:27-35(1982).
CC -1- CATALYTIC ACTIVITY: L-THREONINE + H(2)O -> 2-OXOBUTANOATE + NH(3)
CC + H(2)O.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- ENZYME REGULATION: EACH PROTEIN MOLECULE CAN BIND UP TO FOUR
CC MOLECULES OF AMP, WHICH ACTS AS AN ALLOSTERIC ACTIVATOR TO THE
CC ENZYME. THE ENZYME IS ALSO INHIBITED BY ALPHA-KETO ACIDS AND OTHER
CC CATABOLITES. IT IS ALLOSTERICALLY INHIBITED BY ISOLEUCINE
CC AND ALLOSTERICALLY ACTIVATED BY VALINE.
CC -1- PATHWAY: THREONINE CATABOLISM.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: TO THREONINE DEHYDRATASE BIOSYNTHETIC AND SERINE
CC DEHYDRATASE.
DR PIR: B22317; B22317.
DR STGENE: SG10380; TDCB.
DR INTERPRO: IPR000634; -.
DR PROSITE: PS00165; DEHYDRATASE_SER_THR; PARTIAL.
KW Lyase; Pyridoxal phosphate; Allosteric enzyme.
FT NON_TER 24
FT 24
SQ SEQUENCE 24 AA; 2710 MM; 57BE403BF2C72AF3 CRC64;

Query Match 25.5%; Score 25; DB 1; Length 24;
Best Local Similarity 71.4%; Pred. No. 8.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 14 ILEAKK 20
DB 14 ILEAKK 20

RESULT 6
T2A_PART  STANDARD; PRT; 23 AA.
ID T2A_PART 227173;
AC 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE TRICHOCT MATRIX PROTEIN T2-A (SECRETORY GRANULE PROTEIN T2-A)
DE (TMP 2-A) (FRAGMENT).
GN T2A.
OS Parametium tetraurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Penicillida;
OC Parametium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-D4-2;
RX MEDLINE-96059477; PubMed-7579685;
RA Madeddu L., Gautier M.-C., Vaysse L., Houari A., Sperling L.;
RT "A large multigene family codes for the polypeptides of the
RT crystalline trichocyst matrix in Parametium.";
RT Mol. Biol. Cell 6:649-659(1995).
RN [2]
RP PARTIAL SEQUENCE.
RC STRAIN-D4-2;
RX MEDLINE-95119139; PubMed-7819344;
RA Madeddu L., Gautier M.-C., Je Caer J.P., de Loubresse N., Sperling L.;
RT "Protein processing and morphogenesis of secretory granules in
RT Parametium.";
RT Biochimie 76:329-335(1994).
CC -1- FUNCTION: STRUCTURAL PROTEIN THAT CRYSTALLIZE INSIDE THE
CC TRICHOCT MATRIX.
CC -1- SUBCELLULAR LOCATION: TRICHOCT. THESE ARE ARCHITECTURALLY
CC COMPLEX SECRETORY STORAGE GRANULES-DOCKED AT THE PLASMA MEMBRANE,
CC READY TO RAPIDLY RESPOND TO AN EXOCYTOTIC STIMULUS.
CC -1- SIMILARITY: BELONGS TO THE TMP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U27509; AAA92609.1; -.
CC Polypeptide; Structural protein; Multigene family.
KW Polypeptide; Structural protein; Multigene family.
FT NON_TER 1
FT 1
SQ SEQUENCE 23 AA; 2720 MM; 08972B1F2759BD87 CRC64;

Query Match 24.0%; Score 23.5; DB 1; Length 23;
Best Local Similarity 53.8%; Pred. No. 1.4e+03;
Matches 7; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 8 SDLERRILEAKK 20
DB 11 TDLDRYV-AEOK 22

RESULT 7
T2B_PART  STANDARD; PRT; 23 AA.
ID T2B_PART 027174;
AC 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE TRICHOCT MATRIX PROTEIN T2-B (SECRETORY GRANULE PROTEIN T2-B)
DE (TMP 2-B) (FRAGMENT).
GN T2B.
OS Parametium tetraurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Penicillida;
OC Parametium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-D4-2;
RX MEDLINE-96059477; PubMed-7579685;
RA Madeddu L., Gautier M.-C., Vaysse L., Houari A., Sperling L.;
RT "A large multigene family codes for the polypeptides of the
RT crystalline trichocyst matrix in Parametium.";
RT Mol. Biol. Cell 6:649-659(1995).
RN [2]
RP PARTIAL SEQUENCE.
RC STRAIN-D4-2;
RX MEDLINE-95119139; PubMed-7819344;
RA Madeddu L., Gautier M.-C., Je Caer J.P., de Loubresse N., Sperling L.;
RT "Protein processing and morphogenesis of secretory granules in
RT Parametium.";
RT Biochimie 76:329-335(1994).
CC -1- FUNCTION: STRUCTURAL PROTEIN THAT CRYSTALLIZE INSIDE THE
CC TRICHOCT MATRIX.
CC -1- SUBCELLULAR LOCATION: TRICHOCT. THESE ARE ARCHITECTURALLY
CC COMPLEX SECRETORY STORAGE GRANULES-DOCKED AT THE PLASMA MEMBRANE,
CC READY TO RAPIDLY RESPOND TO AN EXOCYTOTIC STIMULUS.
CC -1- SIMILARITY: BELONGS TO THE TMP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U27510; AAA92610.1; -.
CC Polypeptide; Structural protein; Multigene family.
KW Polypeptide; Structural protein; Multigene family.
FT NON_TER 1
FT 1
SQ SEQUENCE 23 AA; 2706 MM; 184D4B1F2759BD87 CRC64;

```

Query Match 24.0%; Score 23.5; DB 1; Length 23;  
 Best Local Similarity 53.8%; Pred. No. 1.4e+03;  
 Matches 7; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

OY 8 SDELRILEAKX 20  
 DB 11 TLEDRYV-AEK 22

## RESULT 8

CAT2\_FASHE STANDARD; PRT; 14 AA.

AC P80342;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE CATHEPSIN L2 (EC 3.4.22.15) (FRAGMENT).  
 OS Fasciola hepatica (Liver fluke).  
 OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;  
 CC Echinostomida; Echinostomata; Fascioloidae; Fasciolidae; Fasciola.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE-94307282; PubMed-8033913.  
 RA Dowd A.J., Smith A.M., McGonigle S., Dalton J.P.;  
 RT Purification and characterization of a second cathepsin L proteinase  
 RT secreted by the parasitic trematode Fasciola hepatica."  
 RL Eur. J. Biochem. 223:91-98(1994).  
 CC -1- FUNCTION: THIOL PROTEASE THAT ASSISTS THE PARASITE IN BURROWING  
 CC THROUGH THE GUT WALL AND LIVER OF ITS MAMMALIAN HOST.  
 CC -1- SUBUNIT: DIMER OF AN HEAVY AND A LIGHT CHAIN LINKED BY DISULFIDE  
 CC BONDS.  
 CC -1- SUBCELLULAR LOCATION: LYSOSOMAL.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE  
 CC PAPAII FAMILY OF THIOL PROTEASES.  
 DR MEROPS: C01.033; -  
 DR INTERPRO: IPR000169; -  
 DR PROSITE: PS00138; THIOL\_PROTEASE\_CYG; PARTIAL.  
 DR PROSITE: PS00639; THIOL\_PROTEASE\_HIS; PARTIAL.  
 DR PROSITE: PS00640; THIOL\_PROTEASE\_ASN; PARTIAL.  
 KW Hydroxylase; Thiol protease; Lysosome.  
 FT NON\_TER 14 14  
 SQ SEQUENCE 14 AA; 1605 MW; 9CAEAB74E9DA110A CRC64;

OY 6 IPSDLERR 13  
 DB 2 VPKIDRR 9

Query Match 23.5%; Score 23; DB 1; Length 14;  
 Best Local Similarity 37.5%; Pred. No. 9.9e+02;  
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

## RESULT 9

PPBL\_PSEAF STANDARD; PRT; 14 AA.

AC P35482;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE ALKALINE PHOSPHATASE L (EC 3.1.3.1) (L-AP) (FRAGMENT).  
 OS Pseudomonas aeruginosa.  
 CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 CC Pseudomonas.  
 RN [1]  
 RP SEQUENCE.  
 RX STRAIN-H103;  
 RX MEDLINE-93202452; PubMed-8454193;  
 RA Tan A.S.P., Worobec E.A.;  
 RT Isolation and characterization of two immunohistochemically distinct  
 RT alkaline phosphatases from Pseudomonas aeruginosa."  
 RL FEMS Microbiol. Lett. 106:281-286(1993).  
 CC -1- FUNCTION: HAS BOTH A PHOSPHOMONESTERASE AND PHOSPHODIESTERASE

CC ACTIVITY.  
 CC -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN  
 CC ALCOHOL + ORTHOPHOSPHATE (AT A HIGH PH OPTIMUM).  
 CC -1- COFACTOR: METALLOENZYME CONTAINING TWO ZINC ATOMS AND A MAGNESIUM  
 CC ION.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: SECRETED AND TO SOME EXTENT  
 CC PERIPLASMIC.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED 7 HOURS AFTER INTRODUCTION INTO  
 CC PHOSPHATE POOR MEDIA.  
 CC -1- INDUCTION: IN PHOSPHATE POOR MEDIA.  
 CC -1- SIMILARITY: BELONGS TO THE ALKALINE PHOSPHATASE FAMILY.  
 DR INTERPRO: IPR001952; -  
 DR PROSITE: PS00123; ALKALINE\_PHOSPHATASE; PARTIAL.  
 KW Hydroxylase; Zinc; Magnesium; Periplasmic.  
 FT NON\_TER 14 14  
 SQ SEQUENCE 14 AA; 1391 MW; 8B499B7C681B935 CRC64;

Query Match 23.5%; Score 23; DB 1; Length 14;  
 Best Local Similarity 40.0%; Pred. No. 9.9e+02;  
 Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 3 GKMPDLER 12  
 DB 5 GASPLAEIYR 14

## RESULT 10

RL6\_VIBPR STANDARD; PRT; 16 AA.

AC Q56715;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE 50S RIBOSOMAL PROTEIN L6 (FRAGMENT).  
 GN RPL6.  
 OS Vibrio proteolyticus (Aeromonas proteolytica).  
 CC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 15338;  
 RX MEDLINE-97149305; PubMed-8996113;  
 RA Setterquist R.A., Smith G.K., Oakley T.H., Lee Y.H., Fox G.E.;  
 RT "Sequence, overproduction and purification of Vibrio proteolyticus  
 RT ribosomal protein L18 for in vitro and in vivo studies."  
 RL Gene 183:237-242(1996).  
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS  
 CC LOCATED AT THE AMINOACYL-TRNA BINDING SITE OF THE  
 CC PEPTIDYLTRANSFERASE CENTER (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE L6P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: U38943; AB41328.1; -  
 DR INTERPRO: IPR002358; -  
 DR PROSITE: PS00525; RIBOSOMAL\_L6\_1; PARTIAL.  
 KW Ribosomal protein; rRNA-binding.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 16 AA; 1935 MW; ABC19078DF581B6C CRC64;

Query Match 23.5%; Score 23; DB 1; Length 16;  
 Best Local Similarity 46.2%; Pred. No. 1.1e+03;  
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 8 SDELRILEAKX 20

Db 4 ADENVTEKAKR 16

RESULT 11

ID UKAL\_HUMAN STANDARD: PRT; 19 AA.

AC P31940;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 01-JUL-1993 (Rel. 26, Last annotation update)

DE UNKNOWN PROTEIN FROM 2D-PAGE OF EPIDERMAL KERATINOCYTES (SPOT 1118)

DE (FRAGMENT).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

RN [1]

RP SEQUENCE.

RC TISSUE=KERATINOCYTES;

RA MEDLINE=93162043; PubMed=1286667;

RA Rasmussen H.H., Van Damme J., Puye M., Gesser B., Cells J.E.,

RA Vandeckerckhove J.;

RT "Microsequences of 145 proteins recorded in the two-dimensional gel

RT protein database of normal human epidermal keratinocytes.";

RL Electrophoresis 13:960-969(1992).

CC -1- MISCELLANEOUS: ON THE 2D-GEL, THE DETERMINED PI OF THIS UNKNOWN

CC PROTEIN IS: 7.24, ITS MW IS: 23.5 KDA.

DR AARHUS/HEMT-2DPAGE; 1118; IEF.

FT NON\_TER

FT UNSURE

FT NON\_CONS

FT NON\_CONS

FT UNSURE

FT NON\_TER

SO SEQUENCE

19 AA; 2087 MW; EF7515F79D50DE12 CRC64;

Query Match

Best Local Similarity 23.5%; Score 23; DB 1; Length 19;

Matches 3; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 4 KMPSDLERRILEA 17

DB 6 RLTPTEVOEPIITA 19

CONF\_ONCMY

ID COX5LONCMY

STANDARD: PRT; 20 AA.

AC P80329;

DT 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE CYTOCHROME C OXIDASE POLYPEPTIDE VB-HEART (EC 1.9.3.1) (FRAGMENT).

OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;

OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

RN [1]

RP SEQUENCE.

RC TISSUE=HEART;

RA MEDLINE=94237150; PubMed=8181469;

RA Freund R., Kadenbach B.;

RT "Identification of tissue-specific isoforms for subunits Vb and Viba

RT of cytochrome c oxidase isolated from rainbow trout.";

RL Eur. J. Biochem. 221:1111-1116(1994).

DE -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE

DE CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN

DE MITOCHONDRIAL ELECTRON TRANSPORT.

OS -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +

OS 4 FERRIYTOCHROME C.

OC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VB FAMILY.

DR HSSP; P00428; 10CC.

DR INTERPRO; IPR002124; -.

DR PROSITE; PS00848; COX5B; PARTIAL.

KW Oxidoreductase; Inner membrane; Mitochondrion.

FT NON\_TER

FT NON\_TER

SO SEQUENCE

20 AA; 2184 MW; 6A8E00CB33E92BD7 CRC64;

Query Match

Best Local Similarity 23.5%; Score 23; DB 1; Length 20;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 KMPSDLER 12

DB 4 KGIPTEDEQ 12

CONF\_ONCMY

ID PA22\_BOTSC

STANDARD: PRT; 25 AA.

AC P80963;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE PHOSPHOLIPASE A2 HOMOLOG, MYOTOXIN II (FRAGMENT).

OS Bothriechis schlegelii (Eyelash viper).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubridae;

OC Viperidae; Crotalinae; Bothriechis.

RN [1]

RP SEQUENCE.

RC TISSUE=VENOM;

RA MEDLINE=97209510; PubMed=9056257;

RA Arguio Y., Chaves E., Alape A., Rucavado A., Gutierrez J.M.,

RA Lomonte B.;

RT "Isolation and characterization of a myotoxic phospholipase A2 from

RT the venom of the arboreal snake Bothriechis (Bothrops) schlegelii

RL Arch. Biochem. Biophys. 339:260-266(1997).

CC -1- FUNCTION: THIS MYOTOXIC PROTEIN LACKS ENZYMOLOGICAL ACTIVITY.

CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.

DR HSSP; P24605; ICLP.

DR INTERPRO; IPR001211; -.

DR PFAM; PF00068; phoslip; 1.

DR PROSITE; PS00118; PA2\_HIS; PARTIAL.

DR PROSITE; PS00119; PA2\_ASP; PARTIAL.

KW Toxin; Venom; Multigene family.

FT NON\_TER

FT NON\_TER

SO SEQUENCE

25 AA; 2725 MW; C578AB09235EA02A CRC64;

Query Match

Best Local Similarity 23.5%; Score 23; DB 1; Length 25;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ERGXMI 6

DB 4 ELGXMI 9

CONF\_ONCMY

ID CTRB\_THETS

STANDARD: PRT; 22 AA.

AC P81064;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE CYTIN B CHAIN.

OS Theromyzon tessellatum (Leech).

OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea;

OC Rhynchobdellida; Glossiphoniidae; Theromyzon.

RN [1]

RP SEQUENCE.

RX MEDLINE-98055698; PubMed-9395320;  
 RA Chopin V., Biffinger T.V., Stefano G.B., Hatlar I., Salzet M.;  
 RT "Amino-acid-sequence determination and biological activity of cytlh,  
 RT a naturally occurring specific chymotrypsin inhibitor from the leech  
 RT Theromyzon tessatum.",  
 RL Eur. J. Biochem. 249:733-738(1997).  
 CC -1- FUNCTION: INHIBITOR OF CHYMOTRYPSIN.  
 CC -1- SUBUNIT: HETERODIMER OF AN A CHAIN AND A B CHAIN, LINKED BY A  
 CC DISULFIDE BOND.  
 CC -1- SIMILARITY: SOME, TO THE POTATO CHYMOTRYPSIN INHIBITOR I FAMILY  
 CC OF SERINE PROTEASE INHIBITOR.  
 KM Serine protease inhibitor.  
 FT DISULFID 3  
 SQ SEQUENCE 22 AA; 2504 MW; BIC217D5834F97BB CRC64;  
  
 Query Match 22.4%; Score 22; DB 1; Length 22;  
 Best Local Similarity 36.4%; Pred. No. 2.2e+03;  
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
  
 OY 2 RGMIPSDLER 12  
 DB 12 RGEIAKETIER 22  
  
 RESULT 15  
 CR42\_LITCE STANDARD; PRT; 23 AA.  
 AC P56243;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE CAERIN 4.2.  
 OS Litoria caerulea.  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;  
 CC Litoria.  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE-PAROTOID GLAND;  
 RA Stone D.J.M., Maugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;  
 RT "Peptides from Australian frogs. The structures of the caerins from  
 RT Litoria caerulea.",  
 RL J. Chem. Res. 138:910-936(1993).  
 CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE, THAT ADOPTS AN ALPHA HELICAL  
 CC CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN  
 CC DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.  
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL  
 CC GLANDS.  
 CC -1- MASS SPECTROMETRY: MW-2340; METHOD-FAB.  
 KM Antibiotic; Amphibian skin; Amidation.  
 FT MOD\_RES 23  
 SQ SEQUENCE 23 AA; 2343 MW; 83BFD8516ADDC87 CRC64;  
  
 Query Match 22.4%; Score 22; DB 1; Length 23;  
 Best Local Similarity 42.9%; Pred. No. 2.3e+03;  
 Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
  
 OY 4 KMIPSDLERITEA 17  
 DB 7 KSAAGDLASGIVEA 20  
  
 RESULT 16  
 UDP\_LACCA STANDARD; PRT; 23 AA.  
 AC P19662;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE URIDINE PHOSPHORYLASE (EC 2.4.2.3) (UDRPASE) (FRAGMENT).  
 GN UDP.

OS Lactobacillus casei.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;  
 OC Lactobacillus.  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN-ATCC 7469;  
 RX MEDLINE-90381286; PubMed-2119230;  
 RA Avraham Y., Grossowicz N., Yashphe J.;  
 RT "Purification and characterization of uridine and thymidine  
 RT phosphorylase from Lactobacillus casei.",  
 RL Biochim. Biophys. Acta 1040:287-293(1990).  
 CC -1- FUNCTION: THE ENZYMES WHICH CATALYZE THE REVERSIBLE PHOSPHORYLOSIS  
 CC OF PYRIMIDINE NUCLEOSIDES ARE INVOLVED IN THE DEGRADATION OF THESE  
 CC COMPOUNDS AND IN THEIR UTILIZATION AS CARBON AND ENERGY SOURCES,  
 CC OR IN THE RESCUE OF PYRIMIDINE BASES FOR NUCLEOTIDE SYNTHESIS.  
 CC -1- CATALYTIC ACTIVITY: URIDINE + PHOSPHATE -> URACIL + ALPHA-D-RIBOSE  
 CC 1-PHOSPHATE.  
 CC -1- PATHWAY: NUCLEOTIDE AND DEOXYRIBONUCLEOTIDE CATABOLISM.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SIMILARITY: BELONGS TO THE PNP/UDP FAMILY 1 OF PHOSPHORYLASES.  
 DR PIR: S11383; S11383.  
 DR INTERPRO: IPR000845; .  
 DR PROSITE: PS01232; PNP\_UDP\_1; PARTIAL.  
 KM Transferase; Glycosyltransferase.  
 FT UNSURE 1  
 FT OR S.  
 FT NON\_TER 7  
 FT UNSURE 7  
 FT OR T.  
 SQ SEQUENCE 23 AA; 2630 MW; 918B2E2F32P35A17 CRC64;  
  
 Query Match 22.4%; Score 22; DB 1; Length 23;  
 Best Local Similarity 50.0%; Pred. No. 2.3e+03;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
 OY 2 RGMIPSDLE 11  
 DB 1 RLKVIPEKIK 10  
  
 RESULT 17  
 PKR\_VIBCH STANDARD; PRT; 16 AA.  
 ID PKR\_VIBCH  
 AC P96154;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE PHOSPHOGLYCERATE KINASE (EC 2.7.2.3) (FRAGMENT).  
 GN PKR  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OGAMA 395;  
 RX MEDLINE-97136637; PubMed-8982014;  
 RA Carroll P.A., Zhao G., Boyko S.A., Winkler M.E., Calderwood S.B.;  
 RT "Identification, sequencing, and enzymatic activity of the  
 RT erythrose-4-phosphate dehydrogenase gene of Vibrio cholerae.",  
 RL J. Bacteriol. 179:293-296(1997).  
 CC -1- CATALYTIC ACTIVITY: ATP + 3-PHOSPHO-D-GLYCERATE -> ADP +  
 CC 3-PHOSPHO-D-GLYCEROYL PHOSPHATE.  
 CC -1- PATHWAY: SECOND STEP IN THE SECOND PHASE OF GLYCOLYSIS.  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



```

FT  NON_TER 1 1
FT  PEPTIDE 1 23 INTERNAL PEPTIDE VII.
FT  NON_TER 23 23
SQ  SEQUENCE 23 AA: 2689 MW: C84CE139DDDF2B67 CRC64;

Query Match
Best Local Similarity 21.4%; Score 21; DB 1; Length 23;
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 9 DLERRILEAKO 20
    :|:|:|:|
Db 5 EVERKKEAEER 16

RESULT 21
AAMA_BACTR
ID AAMA_BACTR STANDARD; PRT; 24 AA.
AC P37356;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC 3.5.1.14) (L-AMINOACYLASE)
DE (FRAGMENT).
OS Bacillus thermoglucosidasius.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE, AND CHARACTERIZATION.
RA Cho H.-Y., Tanizawa K., Tanaka H., Soda K.;
RT "Thermostable aminocyclase from Bacillus thermoglucosidasius.
RT Purification and characterization."
RL Agric. Biol. Chem. 51:2793-2800(1987).
CC -1- CATALYTIC ACTIVITY: A N-ACYL-L-AMINO ACID + H(2)O -> A FATTY ACID
CC ANION + A L-AMINO ACID.
CC -1- COFACTOR: ACTIVATED BY COBALT (BY SIMILARITY).
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M40, ALSO KNOWN AS THE
CC AHA/HIPO/HYD FAMILY OF HYDROLASES.
KW Hydrolyase; Cobalt.
FT NON_TER 24
SQ SEQUENCE 24 AA: 2846 MW: BC954E4D2B0EC64A CRC64;

Query Match
Best Local Similarity 21.4%; Score 21; DB 1; Length 24;
Matches 4; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 5 MIPSDERRILEAKQ 19
    :|:|:|:|
Db 1 MTNEIKRLVDEYKE 15

RESULT 22
RL41_AGABI
ID RL41_AGABI STANDARD; PRT; 25 AA.
AC P78569;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 60S RIBOSOMAL PROTEIN L41.
GN RL41.
OS Agaricus bisporus (Common mushroom).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Agaricales;
OC Agaricaceae; Agaricus.
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-HORST VI;
RX MEDLINE-97111990; PubMed-8953726;
RA Sonnenberg A.S.M., de Groot P.W.J., Schaap P.J., Baars J.J.P.,
RA Visser J., van Griensven L.J.L.D.;
RT "Isolation of expressed sequence tags of Agaricus bisporus and their
RT assignment to chromosomes."

```

```

RL Appl. Environ. Microbiol. 62:4542-4547(1996).
CC -1- SIMILARITY: BELONGS TO THE L41E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X94764; CA64390.1; -.
DR RIBOSOMAL PROTEIN.
KW RIBOSOMAL PROTEIN.
SQ SEQUENCE 25 AA: 3413 MW: A97629D0143C04EA CRC64;

Query Match
Best Local Similarity 21.4%; Score 21; DB 1; Length 25;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 11 ERRILEAKO 20
    :|:|:|:|
Db 16 KRRKRRARSK 25

RESULT 23
ACPH_RABIT
ID ACPH_RABIT STANDARD; PRT; 6 AA.
AC P25154;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE ACTIAMINO-ACID-RELEASING ENZYME (EC 3.4.19.1) (ACYL-PEPTIDE HYDROLASE)
DE (AHP) (ACTIAMINOACYL-PEPTIDASE) (FRAGMENT).
GN AHPH.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN [1]
RP TISSUE-MUSCLE;
RX MEDLINE-92222120; PubMed-1807161;
RA Krishna R.G., Chin C.C.O., Wold F.;
RT "N-terminal sequence analysis of N alpha-acetylated proteins after
RT unblocking with N-acetylaminocetyl-peptide hydrolase."
RL Anal. Biochem. 199:45-50(1991).
CC -1- FUNCTION: THIS ENZYME CATALYZES THE HYDROLYSIS OF THE AMINO-
CC TERMINAL PEPTIDE BOND OF AN N-ACETYLATED PEPTIDE TO GENERATE
CC AN N-ACETYLATED AA AND A PEPTIDE WITH A FREE AMINO-TERMINUS.
CC IT REPRESENTATIVELY CLEAVES OFF AC-ALA, AC-MET AND AC-SER.
CC -1- CATALYTIC ACTIVITY: ACTIAMINOACYL-PEPTIDE + H(2)O -> ACTIAMINO
CC ACID + PEPTIDE.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9C; ALSO KNOWN AS THE
CC PROLYL OLIGOPEPTIDASE FAMILY.
CC MEROPS: S09.004; -.
DR INTERPRO: IPR002471; -.
DR PROSITE: PS00708; PRO_ENDOPEP_SER; PARTIAL.
KW Hydrolyase; Acetylation.
FT MOD_RES 1 1 ACETYLATION.
FT NON_TER 6 6
SQ SEQUENCE 6 AA: 775 MW: 6732D6C40B16F000 CRC64;

Query Match
Best Local Similarity 20.4%; Score 20; DB 1; Length 6;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 10 IERRIL 15
    :|:|:|:|
Db 1 MEROVL 6

```

## RESULT 24

YATP\_XANCP STANDARD: PRT: 14 AA.  
 ID YATP\_XANCP 003397:  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 01-OCT-1993 (Rel. 27, Last annotation update)  
 DE HYPOTHETICAL PROTEIN IN FRUK 5' REGION (ORF1) (FRAGMENT).  
 GN MTP.  
 OS Xanthomonas campestris (pv. campestris).  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 OC Xanthomonas.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-92011547; Pubmed-1655739;  
 RA de Crecy-Lagard V., Bouvet O.M., Lejeune P., Danchin A.;  
 RT "Fructose catabolism in Xanthomonas campestris pv. campestris.  
 RT Sequence of the PTS operon, characterization of the fructose-specific  
 RT enzymes.";  
 RL J. Biol. Chem. 266:18154-18161(1991).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M69242; AAA27600.1; -.  
 DR PIR: C40944; C40944.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 14 AA; 1585 MW; 72EA9AA3CE16CCDB CRC64;

## Query Match

Best Local Similarity 20.4%; Score 20; DB 1; Length 14;  
 Pred. No. 2.8e+03;

Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 13 RILEAKOK 20  
 | | | | |  
 DB 4 RALEAKRE 11

## RESULT 25

UE15\_HORVU STANDARD: PRT: 15 AA.  
 ID UE15\_HORVU P34938:  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE UNKNOWN ENDOSPERM PROTEIN E-15/E-16/E-17 (FRAGMENT).  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.  
 RN [1]  
 RP SEQUENCE.  
 RP STRAIN-CV. H354-295-2-5; TISSUE-STARCHY ENDOSPERM;  
 RX MEDLINE-94170739; Pubmed-8125056;  
 RA Flengsrud R.;  
 RT "Separation of acidic barley endosperm proteins by two-dimensional  
 RT electrophoresis.";  
 RL Electrophoresis 14:1060-1066(1993).  
 CC -1- TISSUE SPECIFICITY: STARCHY ENDOSPERM.  
 CC -1- DEVELOPMENTAL STAGE: GERMINATION.  
 FT NON\_TER 15  
 SQ SEQUENCE 15 AA; 1735 MW; 266969DDA9D97F CRC64;

## Query Match

Best Local Similarity 20.4%; Score 20; DB 1; Length 15;  
 Pred. No. 3e+03;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 7 PSDLER 12  
 | | | | |  
 DB 8 PYDLNR 13

Search completed: February 5, 2001, 11:34:15  
 Job time: 357 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2001, 11:10:43 ; Search time 141.41 Seconds  
(without alignments)  
9.603 Million cell updates/sec

Title: US-09-343-406-3  
Perfect score: 98  
Sequence: 1 ERGMIPSDLRITLAKOR 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 5001

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: PIR66:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28.6	14	2	G61308	hemocyanin chain 3
2	27.6	24	2	S07645	hsr-omega protein
3	27	27.6	25	H41606	homeotic protein s
4	26.5	11	2	PS0257	31k protein 3208 -
5	26.5	17	2	S61451	hypothetical prote
6	26.5	20	2	H49215	urease (EC 3.5.1.5
7	26.5	20	2	S77983	cytochrome-c oxida
8	25.5	16	2	S35827	uvrX protein - pha
9	25.5	22	2	S35624	uvrX protein - pha
10	25.5	22	2	A53794	tyrosine synthase
11	25.5	22	2	B22317	threonine dehydrat
12	25.5	24	2	B47424	nontoxic type 2 ri
13	24.5	23	2	PS0273	hypothetical prote
14	24.5	16	2	I52226	aldehyde dehydroge
15	24	18	2	PT0239	Ig heavy chain CDR
16	24.5	20	2	S00492	hemocyanin chain I
17	23.5	14	2	S45655	cathepsin L (EC 3.
18	23	16	2	A35552	caldesmon - turkey
19	23.5	16	2	PH1790	T cell receptor al
20	23.5	16	2	E45066	calcium-dependent
21	23	19	2	S11611	ribosomal protein
22	23.5	20	2	S43627	cytochrome-c oxida
23	23.5	20	2	S48746	beta tubulin b-bet
24	23	23	2	S65374	cytochrome-c oxida
25	22.5	23	4	A58505	interleukin-1 beta
26	22.4	8	2	PC4131	hypothetical prote
27	22.4	11	2	S43626	cytochrome-c oxida
28	22.4	13	2	S12388	argA protein - Sal
29	22.4	18	2	B24735	glutathione transf

30	22	22.4	18	2	F49215	urease (EC 3.5.1.5
31	22	22.4	18	2	S46241	NAD(P)-H-flavin oxi
32	22	22.4	20	2	A41717	P100 protein - rat
33	22	22.4	21	2	I40659	hypothetical prote
34	22	22.4	21	2	I40657	hypothetical prote
35	22	22.4	22	2	A39269	LX-1 tumor antigen
36	22	22.4	22	2	S55308	glutathione transf
37	22	22.4	23	2	S16363	opacity protein P.
38	21.5	21.9	18	2	PN0175	glutathione transf
39	21	21.4	11	2	S78765	ribosomal protein
40	21	21.4	12	2	G49215	urease (EC 3.5.1.5
41	21	21.4	14	2	S33601	chaperone, TCP1-re
42	21	21.4	14	2	S68095	calcium-binding pr
43	21	21.4	15	2	A17340	ribonucleoside-dip
44	21	21.4	18	2	P00022	fixA protein - Rhl
45	21	21.4	18	2	A61392	brain-associated s

#### ALIGNMENTS

RESULT 1  
G61308  
hemocyanin chain 3C - Sahara scorpion (fragment)  
C:Species: Androctonus australis (Sahara scorpion)  
C:Date: 17-Jul-1994 #sequence,revision 17-Jul-1994 #text\_change 07-May-1999  
R:Jolles, J.; Jolles, P.; Lamy, J.; Lamy, J.  
FEMS Lett. 106, 289-291, 1979  
A:Title: Structural characterization of seven different subunits in Androctonus austr  
A:Reference number: A61308; MUID:80047238  
A:Accession: G61308  
A>Status: Preliminary  
A:Molecule type: Protein  
A:Residues: 1-14 <JOL>

Query Match 28.6%; Score 28; DB 2; Length 14;  
Best Local Similarity 55.6%; Pred. No. 3.1e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 PSDLRITL 15  
DB 2 PINTORITL 10  
RESULT 2  
S07645  
hsr-omega protein - fruit fly (Drosophila pseudoobscura)  
C:Species: Drosophila pseudoobscura  
C:Date: 20-Feb-1995 #sequence,revision 20-Feb-1995 #text\_change 21-Jul-2000  
C:Accession: S07645  
R:Garbe, J.C.; Bendena, W.G.; Pardue, M.L.  
Genetics 122, 403-415, 1989  
A:Title: Sequence evolution of the Drosophila heat shock locus hsr-omega. I. The nonr  
A:Reference number: S07645; MUID:89357452  
A:Accession: S07645  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-24 <GAR>  
A:Cross-references: EMBL:X16337; NID:g9070; PIDN:CA34395.1; PID:g311729  
C:Genetics:  
A:Gene: FlyBase:HSR93D  
A:Cross-references: FlyBase:FBgn0001234

Query Match 27.6%; Score 27; DB 2; Length 24;  
Best Local Similarity 44.4%; Pred. No. 7.7e+02;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
QY 12 RRIIEAKOR 20  
DB 6 KRIVVAKOR 14

## RESULT 3

H41606 homeotic protein Shox3 - hydromedusa (Sarsia sp.) (fragment)

C:Species: Sarsia sp.

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 24-Sep-1999

C:Accession: H41606

R:Murtha, M.T.; Leckman, J.F.; Ruddle, F.H.

Proc. Natl. Acad. Sci. U.S.A. 88, 10711-10715, 1991

A:Title: Detection of homeobox genes in development and evolution.

A:Reference number: A41606; MUID:92073357

A:Accession: H41606

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-25 &lt;MUR&gt;

A:Cross-references: GB:M81666; NID:g161301; PIDN:AA63528.1; PID:g161302

C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:1-25/Domain: homeobox homology (fragment) &lt;HOX&gt;

Query Match 27.6%; Score 27; DB 2; Length 25;

Best Local Similarity 77.8%; Pred. No. 8.1e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 11 ERRILEAKQ 19

Db 8 ERRILEAKQ 16

## RESULT 4

PS0257

31k protein 3208 - rice (strain Nihonbare) (fragment)

C:Species: Oryza sativa (rice)

C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 11-Apr-1995

C:Accession: PS0257

R:Tsugita, A.; Kano, M.

submitted to JIPID, April 1993

A:Reference number: PS0209

A:Accession: PS0257

A:Molecule type: protein

A:Residues: 1-11 &lt;TSU&gt;

A:Experimental source: callus, strain Nihonbare

C:Comment: molecular weight 31k, pI 5.9.

Query Match 26.5%; Score 26; DB 2; Length 11;

Best Local Similarity 83.3%; Pred. No. 4.9e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 IPSDLE 11

Db 2 IPSDLE 7

## RESULT 5

S61451

hypothetical protein 2 - Campylobacter jejuni (fragment)

C:Species: Campylobacter jejuni

C:Date: 05-Dec-1998 #sequence\_revision 05-Dec-1998 #text\_change 04-Mar-2000

C:Accession: S61451

R:Yao, R.; Burr, D.H.; Doig, P.; Trust, T.J.; Niu, H.; Guerry, P.

Mol. Microbiol. 14, 883-893, 1994

A:Title: Isolation of motile and non-motile insertional mutants of Campylobacter jejuni.

A:Reference number: 140776; MUID:95231295

A:Accession: S61451

A&gt;Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-17 &lt;YAO&gt;

A:Cross-references: GB:U09019; NID:g1177809; PIDN:AA86823.1; PID:g1177811

Query Match 26.5%; Score 26; DB 2; Length 17;  
 Best Local Similarity 46.2%; Pred. No. 7.7e+02;  
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 8 SDLERRILEAKQ 20

Db 2 SDLERRILEAKQ 14

## RESULT 6

H49215

urease (EC 3.5.1.5) small chain Urea - Helicobacter mustelae (fragment)

C:Species: Helicobacter mustelae

C:Date: 19-Dec-1993 #sequence\_revision 25-Apr-1997 #text\_change 30-May-1997

C:Accession: H49215

R:Turnbull, G.R.; Hoj, P.B.; Horne, R.; Mee, B.J.

Infect. Immun. 60, 5259-5266, 1992

A:Title: Purification and characterization of the urease enzymes of Helicobacter spec

A:Reference number: A49215; MUID:93084378

A:Accession: H49215

A&gt;Status: preliminary

A:Molecule type: protein

A:Residues: 1-20 &lt;TUR&gt;

A&gt;Note: sequence extracted from NCBI backbone (NCBIP:119490)

C:Superfamily: urease 26k chain; urease 11k chain homology; urease 12k chain homology

C:Keywords: hydrolase

Query Match 26.5%; Score 26; DB 2; Length 20;

Best Local Similarity 33.3%; Pred. No. 9.1e+02;

Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 KMIPSDLERRIL 15

Db 2 KLTPEKLDKML 13

## RESULT 7

S77983

cytochrome-c oxidase (EC 1.9.3.1) chain Vb - bigeye tuna (fragment)

C:Species: Thunnus obesus (bigeye tuna)

C:Date: 17-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 30-Jan-1998

C:Accession: S77983

R:Arnold, S.; Lee, J.; Kim, M.; Song, E.; Linder, D.; Lottspeich, F.; Kadenbach, B.

submitted to the Protein Sequence Database, June 1997

A:Reference number: S77980

A:Accession: S77983

A:Molecule type: protein

A:Residues: 1-20 &lt;ARN&gt;

A:Experimental source: heart; liver

C:Genetics:

A:Genome: nuclear

C:Function:

A:Pathway: oxidative phosphorylation; respiratory chain

C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membr

Query Match 26.5%; Score 26; DB 2; Length 20;

Best Local Similarity 50.0%; Pred. No. 9.1e+02;

Matches 8; Conservative 1; Mismatches 1; Indels 6; Gaps 1;

QY 4 KMIPSD-----LERR 13

Db 4 KGIPTDEQATGLERR 19

## RESULT 8

S35627

uvrX protein - phage T6 (fragment)

C:Species: phage T6

C:Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 08-Oct-1999

C:Accession: S35627

R:Winkler, M.; Rueger, W.

Nucleic Acids Res. 21, 1500, 1993  
 A>Title: Cloning and sequencing of the genes of beta-glucosyl-HMC-alpha-glucosyl-transferase  
 A:Reference number: S35622; MID:93219141  
 A:Accession: S35627  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-16 <MIN>  
 A:Cross-references: EMBL:G68725; NID:G296439; PID:CAA48668.1; PID:G296441  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992

Query Match 25.5%; Score 25; DB 2; Length 16;  
 Best Local Similarity 40.0%; Pred. No. 1e+03;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 8 SDELRILEA 17  
 :||:|:|:|  
 DB 4 ADKSRILKA 13

RESULT 9  
 S35624  
 uvex protein - phage T2 (fragment)

C:Species: phage T2  
 C>Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 08-Oct-1999

C:Accession: S35624

R:Winkler, M.; Rueger, W.

Nucleic Acids Res. 21, 1500, 1993

A>Title: Cloning and sequencing of the genes of beta-glucosyl-HMC-alpha-glucosyl-transferase

A:Reference number: S35622; MID:93219141

A:Accession: S35624

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-22 <MIN>

A:Cross-references: EMBL:G68724; NID:G296435; PID:CAA48665.1; PID:G296438

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992

Query Match 25.5%; Score 25; DB 2; Length 22;  
 Best Local Similarity 40.0%; Pred. No. 1.4e+03;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 8 SDELRILEA 17  
 :||:|:|:|  
 DB 4 ADKSRILKA 13

RESULT 10  
 A53794  
 tyrocidine synthase I - Bacillus brevis (fragments)

C:Species: Bacillus brevis

C>Date: 01-Dec-1995 #sequence\_revision 01-Dec-1995 #text\_change 17-Mar-1999

C:Accession: A53794

J. Pavella-Vranci, M.; Pfeiffer, E.; Schroeder, W.; von Doehran, H.; Kleinhauf, H.

J. Biol. Chem. 269, 14962-14966, 1994

A>Title: Identification of the AHP binding site in tyrocidine synthetase 1 by selective

A:Reference number: A53794; MID:9423030

A:Accession: A53794

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-22 <PAV>

Query Match 25.5%; Score 25; DB 2; Length 22;  
 Best Local Similarity 44.4%; Pred. No. 1.4e+03;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 GKMPISLE 11  
 :|:|:|  
 DB 3 GAYVPIDIE 11

RESULT 11

B22317  
 threonine dehydratase (EC 4.2.1.16) - Salmonella typhimurium (fragment)

C:Species: Salmonella typhimurium

C>Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 22-Nov-1996

C:Accession: B22317

R:Kim, S.S.; Datta, P.

Biochim. Biophys. Acta 706, 27-35, 1982

A>Title: Chemical characterization of biodegradative threonine dehydratases from two

A:Reference number: A80645; MID:83023208

A:Accession: B22317

A:Molecule type: protein

A:Residues: 1-24 <KIM>

C:Superfamily: threonine dehydratase

C:Keywords: carbon-oxygen lyase; hydro-lyase; threonine catabolism

Query Match 25.5%; Score 25; DB 2; Length 24;  
 Best Local Similarity 71.4%; Pred. No. 1.6e+03;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 14 ILEAKOR 20  
 :|||:|:|:|  
 DB 14 ILEAKOR 20

RESULT 12

B47424  
 nontoxic type 2 ribosome-inactivating protein ebulin 1 B chain - Sambucus ebulus (fra

C:Species: Sambucus ebulus

C>Date: 16-Feb-1994 #sequence\_revision 25-Apr-1997 #text\_change 23-May-1997

C:Accession: B47424

R:Giropes, T.; Citorres, L.; Iglesias, R.; Ferreras, J.M.; Munoz, R.; Rojo, M.A.; Arias

J. Biol. Chem. 268, 18195-18199, 1993

A>Title: Ebulin 1, a nontoxic novel type 2 ribosome-inactivating protein from Sambucus

A:Reference number: A47424; MID:93352642

A:Contents: leaves

A:Accession: B47424

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-25 <GIR>

A>Note: sequence extracted from NCBI backbone (NCBI:136443)

Query Match 25.5%; Score 25; DB 2; Length 25;  
 Best Local Similarity 50.0%; Pred. No. 1.6e+03;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 6 IPSDLERIL 15  
 :||:|:|:|  
 DB 7 IPAPFTIRILV 16

RESULT 13

PS0273  
 hypothetical protein 3 (repB 3' region) - Thiobacillus ferrooxidans plasmid pTF-FC2 (

C:Species: Thiobacillus ferrooxidans

C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 18-Jun-1993

C:Accession: PS0273

R:Dorington, R.A.; Barden, S.; Rawlings, D.E.

Gene 108, 7-14, 1991

A>Title: The broad-host-range plasmid pTF-FC2 requires a primase-like protein for aut

A:Reference number: J50637; MID:92104508

A:Accession: PS0273

A:Molecule type: DNA

A:Residues: 1-23 <DOR>

A:Cross-references: GB:M64981

C:Genetics:

A:Genome: plasmid

Query Match 25.0%; Score 24.5; DB 2; Length 23;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+03;  
 Matches 6; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 6 IPSDLERRILEA 17  
:|||||  
Db 6 LPAYETRL-LEA 16

## RESULT 14

152226  
aldehyde dehydrogenase (EC 1.-.-.-) - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: 152226  
R:Harper, K.; Jones, D.E.  
Biochem. Biophys. Res. Commun. 152, 940-947, 1988  
A:Title: Characterization of a functional recombinant rat liver aldehyde dehydrogenase:  
A:Reference number: 152226; MUID:88209084  
A:Accession: 152226  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-16 <RES>  
A:Cross-references: GB:M29320; NID:9202853; PIDN:AAA40722.1; PID:9202854  
C:Keywords: oxidoreductase

Query Match 24.5%; Score 24; DB 2; Length 16;  
Best Local Similarity 41.7%; Pred. No. 1.5e+03;  
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 SLDLERRILEAKQ 19  
||:|:|  
Db 5 SDYKVRAREARE 16

## RESULT 15

PT0239  
Ig heavy chain CDR3 region (clone 2-100A) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PT0239  
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and  
A:Reference number: PT0239; MUID:91108337  
A:Accession: PT0239  
A:Molecule type: DNA  
A:Residues: 1-18 <YAM>  
A:Experimental source: B lymphocyte  
C:Keywords: heterotetramer; immunoglobulin

Query Match 24.5%; Score 24; DB 2; Length 18;  
Best Local Similarity 71.4%; Pred. No. 1.7e+03;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 LERRILE 16  
|:|:|  
Db 6 LORRALE 12

## RESULT 16

S00492  
hemocyanin chain Ia - Japanese spiny lobster (fragment)  
C:Species: Panulirus japonicus (Japanese spiny lobster)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 01-Aug-1997  
C:Accession: S00492  
R:Makino, N.; Kimura, S.  
Eur. J. Biochem. 173, 423-430, 1988  
A:Title: Subunits of Panulirus japonicus hemocyanin. 1. Isolation and properties.  
A:Reference number: S00492; MUID:88196131  
A:Accession: S00492  
A:Molecule type: protein  
A:Residues: 1-20 <MAK>  
C:Superfamily: hemocyanin

C:Keywords: copper; hemolymph; hexamer; oxygen carrier

Query Match 24.5%; Score 24; DB 2; Length 20;  
Best Local Similarity 38.5%; Pred. No. 1.8e+03;  
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 7 PSDLERRILEAKQ 19  
|:|:|:|  
Db 5 PGDSTDKLLAQKQ 17

## RESULT 17

S45655  
cathepsin L (EC 3.4.22.15) 2 - liver fluke (fragment)  
N:Alternate names: cysteine proteinase  
C:Species: Fasciola hepatica (liver fluke)  
C:Date: 10-Dec-1994 #sequence\_revision 26-Jul-1996 #text\_change 26-Jul-1996  
C:Accession: S45655  
R:Dowd, A.J.; Smith, A.M.; McConigle, S.; Dalton, J.P.  
Eur. J. Biochem. 223, 91-98, 1994  
A:Title: Purification and characterisation of a second cathepsin L proteinase secrete  
A:Reference number: S45655; MUID:94307282  
A:Accession: S45655  
A:Molecule type: protein  
A:Residues: 1-14 <DOM>  
C:Keywords: cysteine proteinase; hydrolase

Query Match 23.5%; Score 23; DB 2; Length 14;  
Best Local Similarity 37.5%; Pred. No. 1.8e+03;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 IPSDLERR 13  
:|:|:|  
Db 2 VPKDKRR 9

## RESULT 18

A35552  
caldesmon - turkey (fragment)  
C:Species: Meleagris gallopavo (common turkey)  
C:Date: 21-Sep-1990 #sequence\_revision 21-Sep-1990 #text\_change 24-Jun-1993  
C:Accession: A35552  
R:Velaz, L.; Ingraham, R.H.; Chalovich, J.M.  
J. Biol. Chem. 265, 2929-2934, 1990  
A:Title: Dissociation of the effect of caldesmon on the ATPase activity and on the bi  
A:Reference number: A35552; MUID:90153926  
A:Accession: A35552  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-16 <VEL>

Query Match 23.5%; Score 23; DB 2; Length 16;  
Best Local Similarity 50.0%; Pred. No. 2.1e+03;  
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 9 DLERRILEAKQ 20  
|:|:|:|  
Db 2 DFERRRLRQK 13

## RESULT 19

PH1790  
T cell receptor alpha chain V region (clone 2PBL V alpha 24-6) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Jul-1999  
C:Accession: PH1790  
R:Porcelli, S.; Fockey, C.E.; Brenner, M.B.; Balk, S.P.  
J. Exp. Med. 178, 1-16, 1993  
A:Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blo

A:Reference number: PH1754; MUID:93301585  
 A:Accession: PH1790  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-16 <PDB>

Query Match 23.5%; Score 23; DB 2; Length 16;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+03;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ERKMP8 8  
 DB 5 ERKMP8 12

## RESULT 20

E45066  
 calcium-dependent carbohydrate-binding protein, L-34b - bovine (fragment)  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 23-Mar-1995  
 C:Accession: E45066  
 R:Kojima, K.; Ogawa, H.K.; Sano, N.; Yamamoto, K.; Irimura, T.; Osawa, T.; Matsumoto, I.  
 J. Biol. Chem. 267, 20535-20539, 1992  
 A:Title: Carbohydrate-binding proteins in bovine kidney have consensus amino acid sequen  
 A:Reference number: A45066; MUID:93015942  
 A:Accession: E45066  
 A:Molecule type: protein  
 A:Residues: 1-16 <KOJ>  
 A:Experimental source: kidney  
 A:Note: sequence extracted from NCBI backbone (NCBIP:116212)

Query Match 23.5%; Score 23; DB 2; Length 16;  
 Best Local Similarity 71.4%; Pred. No. 2.1e+03;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 GKMP8D 9  
 DB 3 GKMP8D 9

## RESULT 21

S11611  
 ribosomal protein S2 [validated] - Halobacterium salinarum (fragment)  
 N:Alternate names: ribosomal protein HS5A  
 C:Species: Halobacterium salinarum  
 C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 21-Jul-2000  
 C:Accession: S11611  
 R:Raguch, M.; Visentin, L.P.; Zuker, M.; Matheson, A.T.; Roy, C.; Strom, A.R.  
 Zbl. Bakt. Hyg. I. Abt. Orig. C 3, 200-208, 1982  
 A:Title: Amino-terminal sequences of ribosomal proteins from the 30S subunit of archaea  
 A:Reference number: S11609  
 A:Accession: S11611  
 A:Molecule type: protein  
 A:Residues: 1-19 <RAG>  
 A:Note: the protein is designated as ribosomal protein HS5A  
 C:Keywords: protein biosynthesis; ribosome

Query Match 23.5%; Score 23; DB 2; Length 19;  
 Best Local Similarity 44.4%; Pred. No. 2.5e+03;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 8 SDLER16 16  
 DB 5 SDLER16 13

## RESULT 22

S43627  
 cytochrome-c oxidase (EC 1.9.3.1) chain Vb-H - trout (fragment)  
 C:Species: Salmo sp. (trout)

C:Date: 19-Mar-1997 #sequence\_revision 01-Aug-1997 #text\_change 02-Jul-1998  
 C:Accession: S43627  
 R:Freund, R.; Kadenbach, B.  
 Eur. J. Biochem. 221, 1111-1116, 1994  
 A:Title: Identification of tissue-specific isoforms for subunits Vb and VIIa of cyto  
 A:Reference number: S43624; MUID:94237150  
 A:Accession: S43627  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-20 <FRE>  
 C:Keywords: electron transfer; membrane-associated complex; oxidoreductase; respirato

Query Match 23.5%; Score 23; DB 2; Length 20;  
 Best Local Similarity 55.6%; Pred. No. 2.6e+03;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 4 KMP8DLR 12  
 DB 4 KMP8DLR 12

## RESULT 23

S48746  
 beta tubulin b-beta-2 - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 17-Mar-1999  
 C:Accession: S48746  
 R:Mary, J.; Redeker, V.; Le Caer, J.P.; Prome, J.C.; Rossier, J.  
 FEBS Lett. 353, 89-94, 1994  
 A:Title: Class I and IV beta-tubulin isoforms expressed in adult mouse brain are glu  
 A:Reference number: S48741; MUID:95010772  
 A:Accession: S48746  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-20 <MAR>

Query Match 23.5%; Score 23; DB 2; Length 20;  
 Best Local Similarity 45.5%; Pred. No. 2.6e+03;  
 Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 ERKMP8D 11  
 DB 3 ERKMP8D 13

## RESULT 24

S65374  
 cytochrome-c oxidase (EC 1.9.3.1) chain IV, cardiac - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 16-Jul-1999  
 C:Accession: S65374  
 R:Schaeffer, H.; Noack, H.; Halausk, W.; Brandt, U.; von Jagow, G.  
 Eur. J. Biochem. 230, 235-241, 1995  
 A:Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-t  
 A:Reference number: S65372; MUID:95324529  
 A:Accession: S65374  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-23 <SCH>  
 C:Superfamily: cytochrome-c oxidase chain IV  
 C:Keywords: cardiac muscle; heart; oxidoreductase

Query Match 23.5%; Score 23; DB 2; Length 23;  
 Best Local Similarity 50.0%; Pred. No. 3e+03;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 6 IPSDERR 13  
 DB 13 IPSDERR 20

## RESULT 25

A58505

Interleukin-1 beta - synthetic (fragment)

C;Species: synthetic

A;Note: Mus musculus (house mouse) gene engineered and expressed in Escherichia coli

C;Date: 08-Jun-1990 #sequence\_revision 13-Feb-1997 #text\_change 13-Feb-1997

C;Accession: A58505

R;Huang, J.J.; Newton, R.C.; Rutledge, S.J.; Horuk, R.; Matthew, J.B.; Covington, M.; Li

J.; Immunol. 140, 3838-3843, 1988

A;Title: Characterization of murine IL-1-beta. Isolation, expression, and purification.

A;Reference number: A27627; MUID:88229074

A;Accession: A58505

A;Molecule type: DNA

A;Residues: 1-23 &lt;HUA&gt;

A;Cross-references: GB:M20874; NID:g208638; PID:g208639

F;2-23/Region: Interleukin-1 beta (fragment)

Query Match 23.0%; Score 22.5; DB 4; Length 23;

Best Local Similarity 35.3%; Pred. No. 3.6e+03;

Matches 6; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 5 MIP-SDLERRILEAKQK 20

Db 1 MVIPIQLHYRLRDEQOK 17

Search completed: February 5, 2001, 11:10:45  
Job time: 1056 sec